Costly Self Control and Random Self Indulgence[1]

Eddie Dekel[2] Barton L. Lipman[3]

First Draft
May 2010

Current Draft
November 2011

[1]We thank Larry Epstein, Faruk Gul, Jawwad Noor, Andy Postlewaite, Todd Sarver, Wolfgang Pesendorfer, three anonymous referees and numerous seminar audiences for helpful comments. We also thank the National Science Foundation, grants SES–0820333 (Dekel) and SES–0851590 (Lipman), for support for this research.

[2]Economics Dept., Northwestern University, and School of Economics, Tel Aviv University E–mail: dekel@nwu.edu.

[3]Boston University. E–mail: blipman@bu.edu.

Abstract

We study the *random Strotz* model, a version of the Strotz (1955) model with uncertainty about the nature of the temptation that will strike. We show that the random Strotz representation is unique and characterize a comparative notion of "more temptation averse." Also, we demonstrate an unexpected connection between the random Strotz model and a generalization of the Gul–Pesendorfer (GP) (2001) model of temptation which allows for the temptation to be uncertain and which we call random GP. In particular, a preference over menus has a random GP representation iff it also has a representation via a random Strotz model with sufficiently smooth uncertainty about the intensity of temptation. We also show that choices of menus combined with choices from menus can distinguish the random GP and random Strotz models.

1 Introduction

In this paper, we explore the *random Strotz* model which is a version of the classic Strotz (1955) model of temptation that adds uncertainty about the nature of the temptation. Uncertainty is both a plausible and a useful hypothesis regarding temptation. Such uncertainty is frequently a key part of applications of the Strotz model, such as in Battaglini, Benabou, and Tirole (2005), Benabou and Tirole (2004, 2010), Eliaz and Spiegler (2006), and Harris and Laibson (2008). Also, as noted by Caplin and Leahy (2006), for example, uncertainty can "smooth out" the discontinuities present in the usual nonstochastic Strotz model.

The resulting model has some surprising connections to more recent models of temptation. In particular, we show a sense in which the random Strotz model nests the self–control model of Gul and Pesendorfer (2001) (henceforth GP). Specifically, the latter exhibits the same commitment behavior — that is, the same choice of "menus" from which future choices will be made — as a particular class of random Strotz models. In addition, a random generalization of the GP model is equivalent in this sense to the class of all randomizations over Strotz models with sufficiently smooth uncertainty about the intensity of temptation (in a sense to be made precise).

We also show that commitment behavior is sufficient to identify the random Strotz model uniquely. In other words, any commitment behavior is consistent with at most one random Strotz model. Given that the random Strotz model is uniquely identified from such behavior, we can characterize how commitment choices vary as we change aspects of the representation. More specifically, we show that a certain kind of first–order stochastic dominance shift upward in the intensity of temptation faced corresponds to an increase in the agent's concern about temptation.[1]

Our results are useful for several reasons. First, they clarify the foundations of the random Strotz model and some of its properties. For example, our uniqueness and comparative results should be valuable in the study of temptation with uncertainty as in the papers cited above. Also, the relationship between the random Strotz and random GP models yields a simple axiomatization of a subclass of random Strotz models. Specifically, Stovall (2010) gives an axiomatic characterization of the random GP model for the case where the support of the measure is finite. As we explain in more detail in Section 3, the connection between random Strotz and random GP that we demonstrate thus implies that Stovall's axioms characterize a subclass of random Strotz models.

[1]By contrast, the random GP model is not identified in the same way as random Strotz. When commitment behavior is consistent with at least one random (or nonrandom) GP model, it is consistent with infinitely many distinct random GP models as we show in Section 4.

Second, the connection between the random Strotz and GP models has significant methodological implications. Most of the work on temptation has focused on using commitment behavior alone as a means of identifying a model, implicitly or explicitly assuming that subsequent choices from menus can be deduced from commitment behavior.[2] Since we show that two very natural models are consistent with the same commitment behavior but predict different choices from menus, such assumptions are not warranted in general. Instead, our results suggest that we should broaden the set of data considered. In particular, if we consider both commitment choices and choices from menus, then we can separate the two models.

Finally, these results may be helpful in other areas where the random Strotz model can be applied. For example, Olszewski (2007) and Ahn (2007) consider models of ambiguity where an act is viewed not as a function from states to consequences but as a set of lotteries, where this is interpreted as a set of consequences. (See also related work by Gajdos, Hayashi, Tallon, and Vergnaud (2008).)[3] In other words, a menu is interpreted not as a set of options that the agent will choose from later but as a set of possible outcomes from which "Nature" will choose later. Under this interpretation, the random Strotz model represents the agent as forming various theories about what guides Nature's choices. Similarly, any model of control rights necessarily has a Strotzian aspect to it, in that an agent must evaluate his utility from the expected choices by another agent given particular constraints. Our uniqueness and comparative results should be useful for such models. Also, as discussed earlier, our results can be combined with those of Stovall (2010) to provide a way to axiomatize such representations.

The basic point that the GP representation can be rewritten in terms of a random determination of which self has control has been made before, though in very different ways. In particular, Benabou and Pycia (2002) note that the GP representation can be written as the equilibrium payoff of a game between the current and future self engaging in a costly battle for control. Also, Chatterjee and Krishna (2007) show that a preference with a GP representation also has a representation where there is a menu–dependent probability that the choice is made by the tempted self, with the choice made by the untempted self otherwise. Unfortunately, the properties of the function relating menus to probabilities over control make it difficult to interpret in general.[4] Our result provides

[2]See, for example, Gul and Pesendorfer (2001) and Dekel, Lipman, and Rustichini (2009).

[3]The Steiner point, which plays a significant role in the analysis of Gajdos, Hayashi, Tallon, and Vergnaud, provides an interesting connection between their work and random Strotz. One definition of the Steiner point of a set of lotteries is that it is the expected value of the lottery chosen by an expected utility preference which is drawn at random from a uniform distribution. Thus it is the expected choice by a particular random Strotz agent.

[4]The published version of Chatterjee and Krishna's paper, Chatterjee and Krishna (2009), considers only the case where this probability is independent of the menu. While this provides more structure, the constant probability model no longer nests GP. On the other hand, this version of their model is a special case of the random Strotz model we consider.

a tighter connection through a model which is an interesting and natural alternative formulation in its own right.

While not the main purpose of their work, Fudenberg and Levine (2006, 2010a, 2010b)'s dual–selves model also gives a connection between GP and multiple–selves models. Our approach enables us to show an unexpected connection between the Fudenberg–Levine model and random Strotz. Roughly speaking, an adaptation of our result linking GP and random Strotz shows that we can recast a version of the dual–selves model which is discussed by Noor and Takeoka (2010b) as random Strotz model. See Section 3 for a more precise statement.

The next section defines the model and the representations considered. In Section 3, we relate random Strotz representations to (random) GP representations. Section 4 shows the uniqueness and comparative results described above. In Section 5, we discuss choice from menus. Proofs not contained in the text are in the Appendix or supplementary appendix.

2 Definitions

Fix a finite set Z of "prizes" or outcomes, let $\Delta(Z)$ denote the set of lotteries over Z, and let X denote the set of *menus*, *i.e.*, the set of compact and nonempty subsets of $\Delta(Z)$. The current self has a preference over X, denoted $\succeq$, interpreted as a preference regarding how much commitment to impose on subsequent choices. (In Section 5, we discuss choices from menus.) Throughout, we assume that $\succeq$ is *nontrivial* in the sense that there exist $x, y \in X$ such that $x \succ y$.

A function $w : \Delta(Z) \to \mathbf{R}$ is an *expected utility (EU) function* if $w(\lambda\alpha + (1-\lambda)\beta) = \lambda w(\alpha) + (1-\lambda)w(\beta)$ for all $\lambda \in [0,1]$ and $\alpha, \beta \in \Delta(Z)$. Both the Strotz and GP representations use two expected utility functions, $u, v : \Delta(Z) \to \mathbf{R}$. The Strotz representation uses u and v to evaluate a menu x by

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

where $B_v(x)$ is the set of best elements of x according to v. That is,

$$B_v(x) = \{\beta \in x \mid v(\beta) \geq v(\alpha), \quad \forall \alpha \in x\}.$$

Intuitively, v represents the preference of the future self who will be completely self indulgent, choosing from the menu as he wishes, breaking ties in favor of the current self who has utility function u.

One unfortunate feature of the Strotz model is that the agent's utility depends discontinuously on the commitments he makes. This occurs because when the choosing self is almost indifferent, the current self may still have strong preferences regarding the choices. A small change in commitments can then create indifference for the chooser, leading to use of the tie–breaking rule and a large change in the current self's payoff. Hence small changes in commitments can have big effects on the current self's payoff.[5] This discontinuity is both intuitively implausible and analytically inconvenient.

The representation introduced by GP is continuous and hence avoids this problem. We say that a GP representation is a pair (u, v) such that a menu x is evaluated by the function

$$V_{GP}(x) = \max_{\beta \in x}[u(\beta) + v(\beta)] - \max_{\alpha \in x} v(\alpha).$$

This representation also has an interesting interpretation. As GP emphasize, the agent chooses from the menu the item which maximizes $u + v$, not v. In this sense, the GP model seems behaviorally richer than the Strotz model as the agent shows partial self control by compromising between u and v instead of simply maximizing v. The term $[\max_{\alpha \in x} v(\alpha)] - v(\beta)$ can be interpreted as the cost of resisting temptation by choosing β instead of maximizing v.

As noted, we consider random versions of the GP and Strotz models. Letting K denote the number of elements of Z, we identify the set of EU functions with $\mathbf{R}^K$ since for any such function, we only need to specify the payoffs to the pure outcomes. We use the Borel σ–algebra over $\mathbf{R}^K$. We say that an EU function is *trivial* if it is a scalar times a vector of 1's. We say a measure μ over $\mathbf{R}^K$ is nontrivial if it assigns zero measure to the set of trivial EU functions.

Definition 1. *A* random Strotz representation *of $\succeq$ is a pair (u, μ) such that u is a nontrivial expected utility function and μ is a nontrivial measure over expected utility functions such that the function*

$$V_{RS}(x) = \int_{\mathbf{R}^K} \max_{\beta \in B_w(x)} u(\beta)\mu(dw)$$

represents the preference.

This is the Strotz representation but where the agent is not sure what his future self's preference will be. It seems quite natural to suppose that an agent may not know exactly what will tempt him in the future or to what extent. Adding uncertainty to the Strotz model also has the potential to resolve the continuity problems noted above since suitably atomless noise ensures that the probability the chooser is indifferent will be zero.

[5] This difficulty is not eliminated by changing the tie–breaking rule.

Consequently, as Caplin and Leahy (2006) show, such atomlessness can ensure existence of an optimal policy in Strotz's sense.

A random GP representation generalizes the notion of a GP representation in a fashion exactly parallel to the above: specifically, the u is fixed but there is a probability measure over the "temptations."

Definition 2. *A* random GP representation *is a pair (u, ν) such that u is a nontrivial expected utility function and ν is a nontrivial measure over expected utility functions such that the function*

$$V_{RGP}(x) = \int_{\mathbf{R}^K} \left\{ \max_{\alpha \in x} [u(\alpha) + v(\alpha)] - \max_{\alpha \in x} v(\alpha) \right\} \nu(dv)$$

represents the preference.

For both random Strotz and random GP, the nontriviality of the measure is without loss of generality in the sense that if a representation exists, then one with a nontrivial measure exists. In both cases, the nontriviality of u is implied by our assumption that $\succeq$ is nontrivial.

3 Random Strotz and Random GP Models

This section discusses the relationship between the models.

Theorem 1. *Any preference with a random GP representation also has a random Strotz representation.*

Proof. We first show the claim for an arbitrary GP representation (u, v). Let W denote the set of expected utility preferences such that $w \in W$ iff there exists $A \in [0, 1]$ with $w = v + Au$. Define a measure μ over W by taking the uniform distribution over A. Let V_{RS} denote the random Strotz representation generated by this measure.

Fix any menu x. Let $\beta^*(A)$ denote any element of x which maximizes u over the set $B_{v+Au}(x)$. Let $\hat{u}(A) = u(\beta^*(A))$ and let $\hat{v}(A) = v(\beta^*(A))$. Note that if multiple elements of x maximize u over $B_{v+Au}(x)$, the values of $\hat{u}(A)$ and $\hat{v}(A)$ do not depend on the particular choice of $\beta^*(A)$. Also, it is easy to show that $\hat{u}$ is nondecreasing in A and hence measurable. Since $\hat{u}$ is also bounded, it is integrable. We have

$$V_{RS}(x) = \int_0^1 u(\beta^*(A))\, dA = \int_0^1 \hat{u}(A)\, dA.$$

Define

$$\mathcal{U}(A) = \hat{v}(A) + A\hat{u}(A) = \max_{\bar{A}\in[0,1]} \hat{v}(\bar{A}) + A\hat{u}(\bar{A}).$$

From the usual argument characterizing incentive compatibility with transferrable utility (see, *e.g.*, Mas–Colell, Whinston, and Green (1995), Proposition 23.D.2, page 888, or Milgrom and Segal (2002), Theorem 2),[6] we have

$$\mathcal{U}(s) = \mathcal{U}(0) + \int_0^s \mathcal{U}'(A)\, dA = \mathcal{U}(0) + \int_0^s \hat{u}(A)\, dA.$$

Hence

$$\mathcal{U}(1) - \mathcal{U}(0) = \int_0^1 \hat{u}(A)\, dA = V_{RS}(x).$$

But $\mathcal{U}(1) = \max_{\beta\in x}[v(\beta) + u(\beta)]$, while $\mathcal{U}(0) = \max_{\beta\in x} v(\beta)$. Hence the left–hand side is the GP representation

To extend to a random GP representation, note that

$$V_{RGP}(x) = \int_{\mathbf{R}^K} \left\{ \max_{\alpha\in x} u(\alpha) + v(\alpha)] - \max_{\alpha\in x} v(\alpha) \right\} \nu(dv)$$

so

$$V_{RGP}(x) = \int_{\mathbf{R}^K} \left\{ \int_0^1 \max_{\beta\in B_{v+Au}(x)} u(\beta)\, dA \right\} \nu(dv) \tag{1}$$

which is a random Strotz representation. ▌

Note the relationship between the random Strotz model constructed in the proof of Theorem 1 and the GP representation (u, v) with the same preference over menus. The random Strotz measure has a distribution over temptations given by taking $v + Au$ where $A \sim U[0, 1]$. Thus the random Strotz has v as the unique "direction" of temptation but has a random intensity defined by the random variable A. For larger A, the choice is based more on the u preference and so the temptation is less intense in this sense. To obtain a random Strotz with the same menu preference as a random GP, we use the probability distribution over directions of temptation given by the random GP and use the same uniform conditional distribution as in equation (1) for the intensity of temptation.

Clearly, not every random Strotz model is also a random GP. A random Strotz model can be discontinuous since random Strotz includes nonrandom Strotz as a special case, while the random GP model inherits the continuity of GP.

[6]To see the connection, consider a standard auction problem or other characterization of incentive compatibility with quasi–linear utility. View A as the type of the agent where this is his valuation for some good. Then $\bar{A}$ plays the role of the agent's report of his type, $\hat{u}(\bar{A})$ is the probability the agent obtains the good if his report is $\bar{A}$, and $\hat{v}(\bar{A})$ is the transfer to him when his report is $\bar{A}$.

So which random Strotz models are also random GP models? Equation (1) gives a partial answer. If we can write the distribution over w's as a distribution over directions (v's) and intensities (A's) of temptation as in equation (1), then the representation is also a random GP model. Writing a distribution over w as a distribution of v and A where $w = v + Au$ is not itself restrictive at all. Obviously, any w can be written as $v + Au$ for appropriately chosen v and A. On the other hand, the requirement that the intensity of temptation A is uniform over $U[0,1]$ independently of the direction of temptation v is quite special.

We show that a much weaker requirement suffices: we can drop the independence of the directions and intensities of temptation and consider any conditional densities over intensities in a large class. More specifically, we generalize the random Strotz representation from the form on the right–hand side of equation (1) to what we call *continuous intensity (CI)* random Strotz representations: those that can be written as

$$\int_{\mathbf{R}^K} \left\{ \int_0^1 \max_{\beta \in B_{v+Au}(x)} u(\beta)\, f(A \mid v)\, dA \right\} \nu(dv)$$

where $f(\cdot \mid v)$ is a lower semicontinuous density.[7]

More precisely, given a random Strotz representation (u, μ), we define a *decomposition* of μ to be a set $\mathcal{V} \subseteq \mathbf{R}^K$, a probability measure $\mu_{\mathcal{V}}$ on $\mathcal{V}$, and a family of conditional probability measures $\mu_A(\cdot \mid v)$ on $\mathbf{R}$ such that for all measurable $E \subseteq \mathbf{R}^K$,[8]

$$\mu(E) = \int_{\mathcal{V}} \mu_A\left(\{A \in \mathbf{R} \mid v + Au \in E\} \mid v\right) \mu_{\mathcal{V}}(dv).$$

We say that (u, μ) is *continuous intensity (CI)* random Strotz representation if there exists a decomposition of μ, say $(\mathcal{V}, \mu_{\mathcal{V}}, \mu_A(\cdot \mid v))$ such that for $\mu_{\mathcal{V}}$–almost all $v \in \mathcal{V}$, μ_A is representable by a lower semicontinuous density. That is, for almost all $v \in \mathcal{V}$ and every measurable $E \subseteq \mathbf{R}$, we have

$$\mu_A(E \mid v) = \int_E f(A \mid v)\, dA$$

where for every $a \geq 0$, $\{A \in \mathbf{R} \mid f(A) > a\}$ is open.

Theorem 2. *The preference $\succeq$ has a random GP representation if and only if it has a CI random Strotz representation.*

[7] This imposes two smoothness conditions. First, the conditional distribution of intensities must have a density. The second condition, lower semicontinuity, seems relatively weak. It strengthens the necessary property of a density that $\{A \mid f(A) > a\}$ be measurable to the requirement that such sets are open. For example, a sufficient condition is that the density be Riemann integrable. (We thank an anonymous referee for pointing this out.) See footnote 10 for comments on the role of the lower semicontinuity requirement.

[8] The measures $\mu_{\mathcal{V}}$ and each $\mu_A(\cdot \mid v)$ are defined on the Borel σ–algebras.

Proof sketch. The proof is in the Appendix. Here we offer an intuition for why the result holds.

First, it is easy to see from equation (1) that if $\succeq$ has a random GP representation, it must have a CI random Strotz representation. Turning to the converse, we focus on the case where there is a single direction of intensity — that is, where $\mathcal{V}$ is a singleton. The extension to the general case is straightforward.

So suppose we have a random Strotz representation of the form

$$V(x) = \int_{\mathcal{A}} \max_{\beta \in B_{v+Au}(x)} u(\beta)\, f(A)\, dA$$

where f is a lower semicontinuous density over the intensity of temptation A and $\mathcal{A}$ is the support of f. We explain how to rewrite V in the form of a random GP representation.

It is easy to see how we can rewrite the random Strotz as a random GP when A is distributed uniformly over some interval other than $[0, 1]$. To see this, simply note that

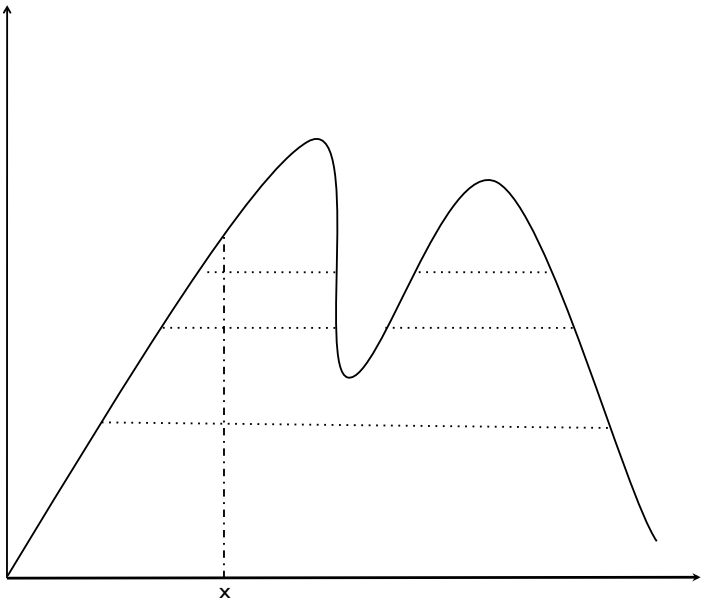
$$\begin{aligned}\int_a^b \max_{\beta \in B_{v+Au}(x)} u(\beta)\, \frac{1}{b-a}\, dA &= \int_0^{b-a} \max_{\beta \in B_{v+au+Au}(x)} u(\beta)\, \frac{1}{b-a}\, dA \\ &= \int_0^{b-a} \max_{\beta \in B_{(v+au+Au)/(b-a)}(x)} u(\beta)\, \frac{1}{b-a}\, dA.\end{aligned}$$

Let $\bar{v} = (v + au)/(b - a)$. Substituting:

$$\int_a^b \max_{\beta \in B_{v+Au}(x)} u(\beta)\, \frac{1}{b-a}\, dA = \int_0^{b-a} \max_{\beta \in B_{\bar{v}+[A/(b-a)]u}(x)} u(\beta)\, \frac{1}{b-a}\, dA = \int_0^1 \max_{\beta \in B_{\bar{v}+\bar{A}u}(x)} u(\beta)\, d\bar{A}$$

where the last equality follows from the change of variables $\bar{A} = A/(b - a)$. From the proof of Theorem 1, we see that this equals the GP representation $(u, \bar{v})$.

From this fact, it is not hard to see that if we can rewrite f as a randomization over uniform distributions over various intervals, then we can rewrite the random Strotz as a randomization over GP representations — that is, as a random GP. It turns out that this can be done if f is lower semicontinuous. To see the idea, consider the density f shown in the figure below.



We identify various uniform distributions by taking the supports to correspond to upper contour sets of f as shown by the dotted lines in the figure. Index these various collections of intervals by a where this denotes the value of $f(x)$ to which they correspond. To match the original density at x, we seek a distribution h such that

$$f(x) = \int_0^{f(x)} \frac{1}{\lambda(\{x' \mid f(x') \geq a\})}\, h(a)\, da$$

where $\lambda(\cdot)$ denotes Lebesgue measure. Note that the problem above has a trivial solution: Let $h(a) = \lambda(\{x' \mid f(x') \geq a\})$. Obviously, this makes the equality hold, so the only question is whether this is a legitimate density. It is easy to see that $h(a) \geq 0$ for all a, so the only issue is whether it integrates to 1. That is, we need to show that

$$\int_0^{f^*} \lambda\left(\{x' \mid f(x') \geq a\}\right)\, da = 1$$

where $f^* = \max_x f(x)$. Note that this integral simply gives another way to compute the area under the density f. For each a between 0 and f^*, it takes the measure of the horizontal line under f at height a and adds these up, getting the area under f which is obviously 1.

For this to complete the proof, we need that the set of x under f at height a is a union of (nontrivial) intervals in order to apply our earlier argument. It is not hard to see that this is true if the density is lower semicontinuous.[9] ∎

Theorems 1 and 2 complement Stovall's (2010) axiomatization of random GP representations. Specifically, Stovall gives axioms on a preference over menus which are necessary and sufficient for the existence of a random GP representation which is finite in the sense that the measure ν has a finite support. From our results, we see that his axioms are also necessary and sufficient for the existence of a CI random Strotz representation which is finite in a certain sense.[10]

Fudenberg and Levine (2006, 2010a, 2010b) and Noor and Takeoka (2010a, 2010b) give nonlinear generalizations of the GP model which also have interesting connections to the random Strotz model. While these papers present a variety of models, we focus on the following class of such nonlinear extensions focused on in Noor and Takeoka (2010b). We define a *nonlinear self–control representation* to be a triple consisting of (u, v, ψ) where u and v are EU functions as before, $\psi : \mathbf{R} \to \mathbf{R}_+$ is an increasing function, and the preference is represented by

$$V_{NSC}(x) = \max_{\beta \in x} \left[u(\beta) - \psi \left(\max_{\alpha \in x} v(\alpha) \right) \left(\max_{\alpha \in x} v(\alpha) - v(\beta) \right) \right].$$

The idea is to model a notion of increasing marginal cost of self–control. More specifically, the self–control costs are scaled up by an increasing function of the "level of temptation" of the menu as measured by $\max_{\alpha \in x} v(\alpha)$. As observed by Lipman and Pesendorfer (2011, footnote 16), an argument similar to the proof of Theorem 1 establishes that V_{NSC} can be written as a different kind of random Strotz: specifically,

$$V_{NSC}(x) = \int_0^{\tau(x)} \max_{\beta \in B_{v+Au}(x)} u(\beta) \frac{1}{\tau(x)} \, dA$$

where $\tau(x) = 1/\psi(\max_{\alpha \in x} v(\alpha))$. Intuitively, this is a random Strotz representation where the level of temptation of the menu affects the distribution over the intensity of temptation. If the menu is more tempting, $\tau(x)$ is smaller, so the A's are more concentrated around zero.

[9] To clarify, the key is that there *exists* a lower semicontinuous density. Obviously, if we have a density which is discontinuous at countably many points, we can choose whether the jumps are "up" or "down" without changing the probability distribution. Hence in such cases, it is without loss of generality to assume the density is lower semicontinuous. Thus the bite of this assumption is only in dealing with densities with uncountably many discontinuities. For such cases, the lower semicontinuity assumption guarantees that upper contour sets are unions of open intervals.

[10] In an earlier version of this paper, Dekel and Lipman (2010) extended Stovall's theorem to give an axiomatization of the class of Lipschitz continuous random Strotz representations. Since there are random GP representations which are not Lipschitz continuous, the class of CI random Strotz representations is still broader than that axiomatized by Dekel and Lipman.

4 Properties of Random Strotz Representations

The relationship between random GP and random Strotz representations yields an easy way to see that random GP representations are not uniquely identified. To be more specific, any preference with a random GP representation has infinitely many such representations — that is, infinitely many probability distributions over "temptations." Furthermore, as illustration of the fact that these representations are nontrivially different, we note that each representation has a distinct prediction regarding choices from menus.[11] To see this most simply, consider a GP representation. From Theorem 1, we know that

$$\max_{\beta\in x}[u(\beta)+v(\beta)]-\max_{\beta\in x}v(\beta)=\int_0^1 \max_{\beta\in B_{v+Au}(x)} u(\beta)\,dA.$$

Fix any partition of $[0,1]$ into N intervals, say $[0,\Delta_1),[\Delta_1,\Delta_2),\ldots,[\Delta_{N-1},1]$. Obviously,

$$\int_0^1 \max_{v+Au} u\,dA=\sum_{n=1}^{N}\int_{\Delta_{n-1}}^{\Delta_n}\max_{v+Au} u\,dA.$$

It is easy to rewrite this as in our proof sketch for Theorem 2 as

$$\sum_{n=1}^{N}\int_0^1 \max_{v_n+Au} u\,[\Delta_n-\Delta_{n-1}]\,dA$$

where

$$v_n=\frac{v+\Delta_{n-1}u}{\Delta_n-\Delta_{n-1}}. \tag{2}$$

Applying Theorem 1 again, this is

$$\sum_{n=1}^{N}[\Delta_n-\Delta_{n-1}]\left\{\max_{\beta\in x}[u(\beta)+v_n(\beta)]-\max_{\beta\in x}v_n(\beta)\right\}, \tag{3}$$

a random GP representation. (One can also show the same conclusion directly by substituting for v_n from equation (2) into equation (3) and rearranging to recover the original GP representation.) As we vary the partition, we vary the collection of v_n's and the corresponding probability distribution. Hence there are infinitely many distinct random GP representations, all of which represent the same preference over menus as the original GP representation.

While all these random GP representations predict the same preferences over menus, they predict very different probability distributions over choices from menus. Specifically,

[11] We thank an anonymous referee for this observation.

as we discuss in more detail in the next section, the random GP above predicts that choice will maximize $u(\beta)+v_n(\beta)$ with probability $\Delta_n-\Delta_{n-1}$. It is easy to see that these choices vary nontrivially with the chosen sequence $\{\Delta_n\}$.

By contrast, if a preference has a random Strotz representation, the representation (the u and the measure over w's) is essentially unique. As an implication, the predicted choice from menus is uniquely identified. To see this, suppose we have a preference $\succeq$ with random Strotz representations (u,μ) and $(\bar{u},\bar{\mu})$. It is easy to see that u and $\bar{u}$ must be the same up to a positive affine transformation. This follows from the fact that both u and $\bar{u}$ must represent the preference over singleton menus. That is, $\{\alpha\} \succeq \{\beta\}$ iff $u(\alpha) \geq u(\beta)$, so $u(\alpha) \geq u(\beta)$ if and only if $\bar{u}(\alpha) \geq \bar{u}(\beta)$. Hence, just as with EU representations, $\bar{u}$ is a positive affine transformation of u.[12]

Similarly, only the choices by a given w matter for the representation, not the level of utility for w from these choices. Thus there is no meaningful distinction between a representation that puts probability p on w from one that puts probability p on $2w$. Hence what is — and what one would naturally want to be — identified is the measure over EU *preferences*, not EU *representations*.

Recall that we identified the space of EU functions with $\mathbf{R}^K$ where K is the number of pure outcomes. Given any Borel set $E \subseteq \mathbf{R}^K$, let

$$\mathcal{B}(E) = \left\{ w \in \mathbf{R}^K \mid aw + b \in E, \text{ some } a \in \mathbf{R}_+, b \in \mathbf{R} \right\}.$$

That is, $\mathcal{B}(E)$ extends E to the set of all EU functions which represent the same preference over lotteries as some function in E. We say that μ and $\bar{\mu}$ are *ordinally equivalent* if for every Borel set $E \subseteq \mathbf{R}^K$, we have $\mu(\mathcal{B}(E)) = \bar{\mu}(\mathcal{B}(E))$. Thus μ and $\bar{\mu}$ are ordinally equivalent if the two measures give the same probability to any given set of EU preferences, ignoring differences between the particular representations chosen for those preferences.

Theorem 3. *If (u,μ) and $(\bar{u},\bar{\mu})$ are random Strotz representations of $\succeq$, then u and $\bar{u}$ are equal up to a positive affine transformation and μ and $\bar{\mu}$ are ordinally equivalent.*

Given that the measure is identified, it is natural to ask how properties of the measure are related to interpretable properties of the preference it represents. To focus on the role of the measure, we compare two preferences with random Strotz representations that have the same u. but different μ's. In particular, the behavioral comparison we consider relates to a version of first–order stochastic dominance (FOSD).[13]

[12] The commitment preference u is also unique up to positive affine transformations in the random GP representation.

[13] Gul and Pesendorfer (2001) also give two comparative notions related to temptation for their model. Their comparatives are very different from ours both in spirit and formally.

We say that $\succeq_2$ is *more temptation averse* than $\succeq_1$ if the restriction of $\succeq_1$ and $\succeq_2$ to singleton menus are the same and if for all menus x and lotteries $\alpha \in x$, whenever $\{\alpha\} \succ_1 x$, we have $\{\alpha\} \succ_2 x$. In other words, whenever $\succeq_1$ strictly prefers committing to a particular choice from the menu rather than leaving the choice open, $\succeq_2$ does as well.[14]

One way to think about this comparative is that it is analogous to comparing the agents in terms of their "willingness to pay" for commitment. To see the idea, note that the more an agent would be willing to give up to achieve commitment, the wider the range of options to which he would prefer to commit. Naturally, in the absence of a common measuring unit such as money, it is difficult to compare two agents in terms of what they are willing to give up to achieve commitment unless they have the same preferences over commitment options. Once we focus on preferences with the same commitment preferences, however, the definition captures a natural notion of greater willingness to pay to avoid temptation.

We now explain how this comparative is reflected in the representation. Suppose we have two preferences with random Strotz representations that can be compared according to this definition. Since they have the same preferences restricted to singleton menus, we can take them to have the same u. So let the random Strotz representation for $\succeq_i$ be denoted (u, μ_i), $i = 1, 2$.

Suppose we have decompositions of μ_1 and μ_2, $(\mathcal{V}_i, \mu^i_{\mathcal{V}_i}, \mu^i_A(\cdot \mid v))$, $i = 1, 2$, that are related in the following way. First, we have $\mathcal{V}_1 = \mathcal{V}_2 \equiv \mathcal{V}$ and $\mu^1_{\mathcal{V}} = \mu^2_{\mathcal{V}} \equiv \mu_{\mathcal{V}}$. Second, for $\mu_{\mathcal{V}}$–almost all $v \in \mathcal{V}$, the conditional distribution $\mu^1_A(\cdot \mid v)$ first order stochastically dominates $\mu^2_A(\cdot \mid v)$. In this case, we say that μ_2 *conditionally–dominates* μ_1.

As discussed earlier, the different v's represent different directions of temptation while the A's measure the intensity of the temptation, where a larger A means less intense temptation. It seems natural that it would be difficult to relate different directions of temptations — e.g., is a temptation to overeat "worse" than a temptation to oversleep? Thus we should require two preferences to be the same regarding the directions of temptation that affect them in order to compare them unambiguously. On the other hand, if one preference has uniformly higher A's in the sense of FOSD, then it has lower intensities and hence has "less trouble" with temptation.

This relationship characterizes our temptation aversion comparison.

Theorem 4. *Fix $\succeq_i$ with random Strotz representation (u, μ_i), $i = 1, 2$. Then $\succeq_2$ is*

[14]This definition is equivalent to one used by Ahn (2007) to compare ambiguity aversion, Sarver (2008) to compare regret attitudes, and Higashi, Hyogo, and Takeoka (2009) to compare aversion to commitment. It is also similar in spirit to the way Epstein (1999) and Ghiradato and Marinacci (2002) define comparisons of ambiguity aversion. Since the random Strotz representation is very different from the representations considered in these papers, their characterization results are quite different as well.

more temptation averse than $\succeq_1$ *if and only if* μ_2 *conditionally–dominates* μ_1.

5 Choice from Menus

To this point, we have focused on the random Strotz and random GP models as representations of preferences over menus. In this sense, we have treated them as models of choice *of* a menu. As we have seen, we cannot, in general, use choice of menus to distinguish the random GP and random Strotz models.

On the other hand, each model also makes predictions about choice *from* menus. In the case of random Strotz, it is natural to interpret the representation (u, μ) as saying that with probability $\mu(w)$, the choice is the one made by w with ties broken in favor of u (where this is stated for measures with finite support for simplicity). In the case of a GP representation (u, v), Gul and Pesendorfer argue that the natural interpretation of the choice from a menu x is that it is some maximizer of $u+v$ from that menu. It is natural to interpret a random GP representation (u, ν) analogously as saying that with probability $\nu(v)$, the choice is that which maximizes $u+v$. In both cases, the representation theorem tells us that we can represent the agent as if these are the choices from menus that he anticipates. If we add to the models the hypothesis that these "anticipated" choices are the actual choices and if we also observe both choices of menus and choices from menus, can we then distinguish random GP and random Strotz?[15,16]

Formally, fix a random Strotz representation (u, μ). We define a *selection function for* (u, μ) to be a measurable function $\beta^* : X \times \text{supp}(\mu) \to \Delta(Z)$ such that $\beta^*(x, w) \in B_u(B_w(x))$ for all $(x, w) \in X \times \text{supp}(\mu)$. That is, the selection function $\beta^*(x, w)$ gives one way that choices could be made from menu x in the random Strotz representation as a function of w. Then we can define a *random choice function* $\rho : X \to \Delta(\Delta(Z))$ by

$$\rho_x(E) = \mu\left(\{w \in \mathcal{W} \mid \beta^*(x, w) \in E\}\right) \tag{4}$$

[15]Unsurprisingly, if we only observe choices from menus, we cannot distinguish these models in general. Both models predict choice from menus in the form of random expected utility, though with a tie–breaking rule in the case of random Strotz. While the tie–breaking distinction can sometimes separate the two models on the basis of choice from menus, this is not generally possible.

[16]The example at the beginning of Section 4 illustrating the nonuniqueness of random GP representations might suggest that the answer is no. One can use the approach we gave to construct a sequence of random GP representations, all with the same preference over menus, whose choice from menus converges to that of a random Strotz representation with the same preference over menus. Thus if this sequence of random GP representations converged to a random GP representation, it would show that the answer to our question is "no." However, one can show that this sequence does not converge to a random GP representation.

for every measurable $E \subseteq x$. We say that such a random choice function is *rationalized* by (u, μ).

Turning to analogous notions for a random GP representation (u, ν), we define a selection function to be a measurable $\hat{\beta}^* : X \times \text{supp}(\nu) \to \Delta(Z)$ such that $\hat{\beta}^*(x, v) \in B_{u+v}(x)$ for all $(x, v) \in X \times \text{supp}(\nu)$. We then define a random choice function ρ generated by this selection function by

$$\rho_x(E) = \nu\left(\left\{v \in \mathbf{R}^K \mid \hat{\beta}^*(x, v) \in E\right\}\right)$$

every measurable $E \subseteq x$. We say such a ρ is rationalized by (u, ν).[17]

Obviously, the case of no temptation is a special case for both models, so if we are to distinguish the models, we must rule out this common case. We say that a preference $\succeq$ over menus *exhibits temptation* if there exist $\alpha, \beta \in \Delta(Z)$ with $\{\alpha\} \succ \{\beta\}$ and $\{\alpha\} \succ \{\alpha, \beta\}$.[18] It's not hard to show that if (u, μ) is a random Strotz representation of $\succeq$, then $\succeq$ exhibits temptation if and only if there is a $w \in \text{supp}(\mu)$ such that w does not represent the same preference over lotteries as u. Similarly, if (u, ν) is a random GP representation of $\succeq$, then $\succeq$ exhibits temptation iff there exists $v \in \text{supp}(\nu)$ such that v does not represent the same preference over lotteries as u.

We show that if a random Strotz representation and a random GP representation generate the same choices from menus, then they cannot represent the same preferences over menus. Specifically, either the commitment preferences (the preferences over singleton menus) will differ or else the random GP agent's preference over menus will be strictly more temptation–averse.[19]

Theorem 5. *Suppose random choice function ρ has both a random Strotz rationalization (u, μ) and a random GP rationalization $(\hat{u}, \nu)$. Let $\succeq_{RS}$ be the preference represented by (u, μ) and $\succeq_{RGP}$ the preference represented by $(\hat{u}, \nu)$. Then if $\hat{u}$ and u represent the same preference over singletons,[20] $\succeq_{RGP}$ is more temptation averse than $\succeq_{RS}$, strictly so if $\succeq_{RGP}$ exhibits temptation.*

[17] These definitions are essentially the same as those used in Gul and Pesendorfer (2006).

[18] One can show that this is equivalent to what GP call $\succeq$ having a preference for commitment if $\succeq$ has either a random Strotz or random GP representation.

[19] A similar argument to the proof of Theorem 5 shows that we could also compare the choices by a random Strotz representation and a random GP representation that correspond to the same preference over menus. Specifically, one can then show that the agent would prefer to commit to the expected behavior under the random GP than to the expected behavior under the random Strotz.

[20] The choice function alone cannot tell us whether the rationalizations have the same preference over singletons or not. It is not hard to show that if a random choice function has a random GP rationalization, we can find such a rationalization with any u function we choose. To see this, consider for simplicity a random GP rationalization (u, ν) where the support of ν is $\{v_1, \ldots, v_J\}$. Fix *any* $\hat{u}$. For $j = 1, \ldots, J$, let $\hat{v}_j = u - \hat{u} + v_j$, so that $\hat{u} + \hat{v}_j = u + v_j$. Define $\hat{\nu}$ by $\hat{\nu}(\hat{v}_j) = \nu(v_j)$. The random GP rationalization (u, ν) says that given menu x, the choice is the one which maximizes $u + v_j$ with probability $\nu(v_j)$. Clearly,

Proof. Without loss of generality, assume $\hat{u} = u$. To show that the comparison holds, we show that $\{\alpha\} \succeq_{RS} x$ implies $\{\alpha\} \succeq_{RGP} x$. To see this, fix any menu x. By definition of the random Strotz representation, we must have $x \sim_{RS} \left\{\int \beta\, \rho_x(d\beta)\right\}$. Hence $\{\alpha\} \succeq_{RS} x$ implies $\{\alpha\} \succeq_{RS} \left\{\int \beta\, \rho_x(d\beta)\right\}$. Since both rationalizations rank singletons the same way, we have $\{\alpha\} \succeq_{RGP} \left\{\int \beta\, \rho_x(d\beta)\right\}$. But, letting $\hat{\beta}^*$ denote a selection function for (u, ν),

$$\begin{aligned}
V_{RGP}(x) &= \int \left\{\max_x [u(\beta) + v(\beta)] - \max_x v(\beta)\right\} \nu(dv) \\
&= \int \left\{u(\hat{\beta}^*(x,v)) + v(\hat{\beta}^*(x,v)) - \max_x v(\beta)\right\} \nu(dv) \\
&\leq \int \left\{u(\hat{\beta}^*(x,v)) + v(\hat{\beta}^*(x,v)) - v(\hat{\beta}^*(x,v))\right\} \nu(dv) \\
&= \int u(\hat{\beta}^*(x,v))\, \nu(dv) = u\left(\int \beta\, \rho_x^{RGP}(d\beta)\right) \\
&= V_{RGP}\left(\left\{\int \beta\, \rho_x^{RGP}(d\beta)\right\}\right).
\end{aligned}$$

Hence $\{\int \beta\, \rho_x(d\beta)\} \succeq_{RGP} x$. So $\{\alpha\} \succeq_{RGP} \left\{\int \beta\, \rho_x(d\beta)\right\} \succeq_{RGP} x$, establishing the desired conclusion.

To see that there must be some menu where the comparison is strict if $\succeq_{RGP}$ exhibits temptation, consider any α and β that satisfy $\{\alpha\} \succ_{RGP} \{\beta\}$ and $\{\alpha\} \succ_{RGP} \{\alpha, \beta\}$. As noted above, this implies that there is $v \in \text{supp}(\nu)$ with $u(\alpha) > u(\beta)$ and $v(\alpha) < v(\beta)$. It is easy to show that this implies that for the menu $x = \{\alpha, \beta\}$, we must have $V_{RGP}(x) < V_{RGP}(\{\int \beta\, \rho_x^{RGP}(d\beta)\})$. Hence the inequality above is strict for such a menu. ∎

To see the intuition behind this result most simply, suppose $\succeq$ has a GP representation and hence also a random Strotz representation. Suppose this preference has $\{\alpha\} \succ \{\alpha, \beta\} \succ \{\beta\}$. In the GP case, this is rationalized by having $u(\alpha) > u(\beta)$, $v(\beta) > v(\alpha)$, and $u(\alpha) + v(\alpha) > u(\beta) + v(\beta)$. These rankings imply that

$$\begin{aligned}
V_{GP}(\{\alpha, \beta\}) &= \max\{u(\alpha) + v(\alpha), u(\beta) + v(\beta)\} - \max\{v(\alpha), v(\beta)\} \\
&= u(\alpha) - [v(\beta) - v(\alpha)].
\end{aligned}$$

Thus the predicted choice is α, the same as the "choice" from the menu $\{\alpha\}$, but the menu is ranked lower than $\{\alpha\}$ because of the self–control cost of $v(\beta) - v(\alpha)$. By contrast, the random Strotz representation would have $V_{RS}(\alpha, \beta) = pu(\alpha) + (1-p)u(\beta)$ for some $p \in (0, 1)$. Thus the random Strotz representation "explains" the fact that $\{\alpha\} \succ \{\alpha, \beta\}$ not by self–control costs but by a nonzero probability of "self–indulgent" behavior under the latter menu.

this is the same thing as saying it is the choice which maximizes $\hat{u} + \hat{v}_j$ with probability $\hat{\nu}(\hat{v}_j)$. Hence $(\hat{u}, \hat{\nu})$ is also a random GP rationalization. In the case of a continuous random Strotz representation, the tie–breaking in favor of u has no effect on choices and so we get a similar indeterminacy.

In other words, the random Strotz model explains the desire for commitment entirely in terms of a fear of succumbing to temptation, while the random GP model explains it in part by this but in part by a desire to avoid self–control costs. Hence if the choices from menus coincide, the self–control costs lead the random GP agent to have a stronger desire for commitment than the random Strotz agent.

A Proof of Theorem 2

As noted in the text, we begin with a random Strotz representation of the form

$$V(x) = \int_{\mathcal{A}} \max_{\beta \in B_{v+Au}(x)} u(\beta)\, f(A)\, dA$$

where f is lower semicontinuous and $\mathcal{A}$ is the support of f. We show how to rewrite V in the form of a random GP representation.

For any $a \geq 0$, let $U(a) = \{A \in \mathbf{R} \mid f(A) > a\}$. Since f is lower semicontinuous, $U(a)$ is open and hence is the union of countably many disjoint open intervals for all a. Let $\mathcal{L}(a)$ denote the Lebesgue measure of $U(a)$. Then we have

$$\begin{aligned} V(x) &= \int_{\mathcal{A}} \max_{\beta \in B_{v+Au}(x)} u(\beta) \left[\int_0^{f(A)} \frac{1}{\mathcal{L}(a)}\, \mathcal{L}(a)\, da \right] dA \\ &= \int_{\mathcal{A}} \int_0^{f(A)} \max_{\beta \in B_{v+Au}(x)} u(\beta) \frac{1}{\mathcal{L}(a)}\, \mathcal{L}(a)\, da\, dA \end{aligned}$$

Let $\bar{f} = \sup_{A \in \mathcal{A}} f(A)$. Then the double integral above is over

$$\left\{(A,a) \in \mathbf{R}^2 \mid A \in \mathcal{A},\ 0 < a < f(A)\right\} = \left\{(A,a) \in \mathbf{R}^2 \mid 0 < a < \bar{f},\ f(A) > a\right\}.$$

Thus the integral is equal to

$$\int_0^{\bar{f}} \left[\int_{U(a)} \max_{\beta \in B_{v+Au}(x)} u(\beta)\, \frac{1}{\mathcal{L}(a)}\, dA \right] \mathcal{L}(a)\, da.$$

Note that $\int_0^{\bar{f}} \mathcal{L}(a)\, da = \int_{\mathcal{A}} f(A)\, dA = 1$. Hence we can view $\mathcal{L}(a)$ as a density over $a \in [0, \bar{f}]$.

Fix any $a \in (0, \bar{f})$. Since $U(a)$ is a union of disjoint open intervals, we can write

$$\int_{U(a)} \max_{\beta \in B_{v+Au}(x)} u(\beta)\, \frac{1}{\mathcal{L}(a)}\, dA = \sum_{k=1}^{\infty} \int_{b_k}^{c_k} \max_{\beta \in B_{v+Au}(x)} u(\beta)\, \frac{1}{\mathcal{L}(a)}\, dA,$$

where (b_k, c_k), $k = 1, 2, \ldots$, is the collection of intervals defining $U(a)$ and we suppress the dependence of the b_k's and c_k's on a for notational simplicity. Rewriting, this is

$$= \sum_{k=1}^{\infty} \frac{c_k - b_k}{\mathcal{L}(a)} \int_{b_k}^{c_k} \max_{\beta \in B_{v+Au}(x)} u(\beta)\, \frac{1}{c_k - b_k}\, dA.$$

As shown in the text,

$$\int_{b_k}^{c_k} \max_{\beta \in B_{v+Au}(x)} u(\beta) \frac{1}{c_k - b_k} \, dA = \max_{\beta \in x}[u(\beta) + v_k(\beta)] - \max_{\beta \in x} v_k(\beta)$$

where $v_k = (v + b_k u)/(c_k - b_k)$. Hence for any a, we have

$$\int_{U(a)} \max_{\beta \in B_{v+Au}(x)} u(\beta) \frac{1}{\mathcal{L}(a)} \, dA = \sum_{k=1}^{\infty} \frac{c_k - b_k}{\mathcal{L}(a)} \left\{ \max_{\beta \in x}[u(\beta) + v_k(\beta)] - \max_{\beta \in x} v_k(\beta) \right\}.$$

Since $\sum_k (c_k - b_k)/\mathcal{L}(a) = 1$, this is a random GP representation. Hence we have established that the random Strotz representation is a randomization over random GP representations and hence is a random GP representation. ▌

B Proof of Theorem 3

We prove this result by showing that if we restrict attention to measures on a particular subspace of $\mathbf{R}^K$ which includes one EU function for each nontrivial EU preference, then the measure is unique. The particular space we use is

$$\mathcal{W} = \{w \in \mathbf{R}^K \mid w \cdot \mathbf{1} = 0, w \cdot w = 1\}$$

where $\mathbf{1}$ is a K vector of 1's. It is easy to see that any nontrivial EU preference is represented by exactly one $w \in \mathcal{W}$.

Theorem 3 is concerned only with the measure of sets of EU functions which are closed under equivalence in the sense that if w is in the set, then every w' equivalent to w is in the set as well. Hence we may as well take our measures to be over $\mathcal{W}$. For the σ–algebra on $\mathcal{W}$, we use the Borel σ–algebra using as our topology on $\mathcal{W}$ the (relativized) usual Euclidean topology on $\mathbf{R}^K$.

The proofs of Lemmas 1 and 2 are straightforward algebra and hence omitted.

Lemma 1. *Fix $w, \bar{w} \in \mathcal{W}$. Then $w \cdot \bar{w} \in [-1, 1]$. Furthermore, $w \cdot \bar{w} = 1$ iff $w = \bar{w}$ and $w \cdot \bar{w} = -1$ iff $w = -\bar{w}$.*

Let $\mathcal{V} = \{v \in \mathcal{W} \mid v \cdot u = 0\}$.

Lemma 2. *For every $w \in \mathcal{W}$, there exists $v \in \mathcal{V}$ and $A \in [-1, 1]$ such that $w = v\sqrt{1 - A^2} + Au$. If $w = u$, then this holds for every $(A, v) \in \{1\} \times \mathcal{V}$, while if $w = -u$, it holds for every $(A, v) \in \{-1\} \times \mathcal{V}$. For every other $w \in \mathcal{W}$, the (A, v) is unique.*

Define an order over $\mathcal{W}$ by $w\ C_u\ \hat{w}$ (read "w is closer to u than $\hat{w}$") if

$$u(\alpha) > u(\beta),\ \hat{w}(\alpha) \geq \hat{w}(\beta)\ \text{ implies }\ w(\alpha) \geq w(\beta).$$

In other words, w is willing to "go along with" u at least as often as $\hat{w}$. Define a set $W \subseteq \mathcal{W}$ to be *closed under* C_u if $w' \in W$ and $w\ C_u\ w'$ implies $w \in W$.

For brevity, let $\eta(A) = \sqrt{1-A^2}$. The proof of the following lemma is in the supplementary appendix.

Lemma 3. *$w_1\ C_u\ w_2$ if and only if there exists $v \in \mathcal{V}$ such that $w_i = v\sqrt{1-A_i^2} + A_i u$, $i = 1, 2$, with $A_1 \geq A_2$.*

Given a function $A^* : \mathcal{V} \to [-1, 1]$, let

$$W(A^*) = \bigcup_{v \in \mathcal{V}} \{w \in \mathcal{W} \mid w = v\sqrt{1-A^2} + Au,\ \text{for some } A \geq A^*(v),\ A \neq -1\}.$$

Note that by excluding $A = -1$, the definition of $W(A^*)$ ensures that $-u \notin W(A^*)$ for any A^*.

Lemma 4. *A set $W \subseteq \mathcal{W}$, $W \neq \mathcal{W}$, is closed and closed under C_u if and only if there exists a lower semi–continuous function A^* such that $W = W(A^*)$ and $A^*(v) > -1$ for all $v \in \mathcal{V}$.*

Proof. Fix any lower semi–continuous function A^* such that $A^*(v) > -1$ for all $v \in \mathcal{V}$. Let $W = W(A^*)$. Since the definition of $W(A^*)$ prevents $-u \in W(A^*)$, $W \neq \mathcal{W}$. From Lemma 3, it is easy to see that W is closed under C_u if and only if $v\sqrt{1-A^2} + Au \in W$ implies $v\sqrt{1-\hat{A}^2} + \hat{A}u \in W$ for all $\hat{A} \in (A, 1]$. Hence the definition of $W(A^*)$ obviously implies W is closed under C_u. Finally, to show that W is closed, fix any sequence w_n converging to w such that $w_n \in W$ for all n. We can write $w_n = \sqrt{1-A_n^2}v_n + A_n u$ for each n. Since $w_n \in W$ for all n, we have $A_n \geq A^*(v_n)$ for all n. Let v denote the limit of v_n and A the limit of A_n. It is easy to see from the proof of Lemma 2 that the v and A associated with a given w depend continuously on w, so we must have $w = \sqrt{1-A^2}v + Au$. Hence $w \in W(A^*)$ if and only if $\lim_{n\to\infty} A^*(v_n) \geq A^*(\lim_{n\to\infty} v_n)$. Since A^* is lower semi–continuous, this holds. Hence W is closed.

For the converse, suppose $W \neq \mathcal{W}$ is closed and closed under C_u. For each v, let

$$A^*(v) = \min\{A \in [-1, 1] \mid v\sqrt{1-A^2} + Au \in W\}.$$

Since W is closed, $A^*(v)$ is well–defined. Since $w\ C_u\ -u$ for all w, the fact that $W \neq \mathcal{W}$ implies $-u \notin W$. Hence $A^*(v) > -1$ for all v. Since W is closed under C_u, for every

$A \geq A^*(v)$, we have $v\sqrt{1-A^2} + Au \in W$, implying that $v\sqrt{1-A^2} + Au \in W$ if and only if $A \geq A^*(v)$. Hence $W = W(A^*)$.

Finally, to see that A^* is lower semi–continuous, again, consider the sequence constructed above. As noted, for any such sequence, $w \in W$ if and only if $\lim_{n\to\infty} A^*(v_n) \geq A^*(\lim_{n\to\infty} v_n)$. Since W is closed, we must have $w \in W$. Hence any jumps in A^* must be downward, so A^* is lower semi–continuous. ∎

We note that if A^* is lower semi–continuous, then it is measurable.

Lemma 5. *Fix any measurable function $A^* : \mathcal{V} \to [-1, 1]$ such that $W(A^*)$ is closed. Then there exists a sequence of positive numbers $\{\varepsilon_n\}$ and a sequence of menus $\{x_n\}$ such that for every random Strotz representation with commitment preference represented by u,*

$$\lim_{n\to\infty} \frac{V(x_n)}{\varepsilon_n} = \mu(W(A^*)).$$

Proof. Fix such an A^* function.

Part 1. First, suppose that A^* is bounded in the sense that $A^*(v)/\sqrt{1-(A^*(v))^2}$ is bounded from above and below.

For each $v \in \mathcal{V}$, let

$$\alpha_v = \frac{1}{K}\mathbf{1} + \varphi v$$

$$\beta_v(\varepsilon) = \alpha_v + \varphi\varepsilon\left[u - \frac{A^*(v)}{a^*(v)}v\right]$$

where $a^*(v) = \sqrt{1-(A^*(v))^2}$. By the boundedness of A^*/a^*, there exists $\varphi > 0$ such that for all sufficiently small $\varepsilon > 0$, every α_v and $\beta_v(\varepsilon)$ is a lottery.

Let $L(v) = \{w \in \mathcal{W} \mid w = av + Au, \text{ some } a \geq 0, A \in -[-1, 1]\}$. Suppose $w = L(v)$ and consider some $\bar{v}$ which may or may not equal v. Then

$$w \cdot \alpha_{\bar{v}} = a\varphi v \cdot \bar{v}$$

while

$$w \cdot \alpha_v = a\varphi.$$

Since $v \cdot \bar{v} \leq 1$, strictly if $\bar{v} \neq v$, we see that $w \cdot \alpha_{\bar{v}} \leq w \cdot \alpha_v$, strictly so for any $\bar{v} \neq v$. Hence if w picks any α, he must pick α_v.

Also, for any v, $u \cdot \alpha_v = 0 < \varphi\varepsilon = u \cdot \beta_v(\varepsilon)$. So u is indifferent among the α's, indifferent among the β's, and prefers the β's to the α's. Hence, letting x_ε denote the

menu consisting of all the α's and β's, we see that $V(x_\varepsilon) = \varphi\varepsilon\mu(W_\varepsilon)$, where

$$W_\varepsilon = \bigcup_{v\in\mathcal{V}}\{w \in L(v) \mid w\cdot\beta_{\bar v}(\varepsilon) \geq w\cdot\alpha_v, \text{ for some } \bar v \in \mathcal{V}\}.$$

We now show that $\lim_{\varepsilon\downarrow 0}\mu(W_\varepsilon) = \mu(W(A^*))$. Note that if $w = av + Au$, then $a = \sqrt{1-A^2}$ and

$$w\cdot\beta_v(\varepsilon) = w\cdot\alpha_v + \varphi\varepsilon\left[A - \frac{A^*(v)}{a^*(v)}a\right],$$

so $w\cdot\beta_v(\varepsilon) \geq w\cdot\alpha_v$ iff $(A/a) \geq (A^*(v)/a^*(v))$. It is not hard to show that this holds iff $A \geq A^*(v)$. Hence for every ε, we have $W(A^*) \subseteq W_\varepsilon$.

Next, we show that if $w \notin W(A^*)$, then there is $\bar\varepsilon > 0$ such that for all $\varepsilon \in (0,\bar\varepsilon)$, we have $w \notin W_\varepsilon$. To show this, suppose not. Then there exists a sequence ε_n converging to zero such that $w \in W_{\varepsilon_n} \setminus W(A^*)$ for all n.

Write $w = av + Au$. Then there exists a sequence $\bar v_n$ such that

$$w\cdot\beta_{\bar v_n}(\varepsilon_n) \geq w\cdot\alpha_v$$

or

$$a\varphi v\cdot\bar v_n + \varphi\varepsilon_n\left[A - \frac{A_n}{a_n}av\cdot\bar v_n\right] \geq a\varphi$$

where $A_n = A^*(\bar v_n)$ and $a_n = \sqrt{1-A_n^2}$. Rearranging yields

$$\varepsilon_n\left[\frac{A}{a} - \frac{A_n}{a_n}v\cdot\bar v_n\right] \geq 1 - v\cdot\bar v_n.$$

Since $v\cdot\bar v_n \leq 1$ and A_n/a_n is bounded from below, we must have $v\cdot\bar v_n \to 1$ as $n\to\infty$. Note for future use that this implies $\bar v_n \to v$. Also, the fact that the right–hand side is nonnegative for all n implies that

$$\frac{A}{a} \geq \frac{A_n}{a_n}v\cdot\bar v_n$$

for all n. Recall that $w \in L(v)$ and $w \notin W^*$. Hence $A < A^*(v)$. So we have

$$\frac{A^*(v)}{a^*(v)} > \frac{A^*(\bar v_n)}{a^*(\bar v_n)}v\cdot\bar v_n.$$

Since $v\cdot\bar v_n \to 1$, we have

$$\frac{A^*(v)}{a^*(v)} \geq \lim_{n\to\infty}\frac{A_n}{a_n},$$

or $A^*(v) \geq \lim_{n\to\infty} A_n$.

By Lemma 4, the fact that $W(A^*)$ is closed that A^* is lower semi–continuous. Hence we have the opposite weak inequality, so $A^*(v) = \lim_{n\to\infty} A_n$.

But recall that

$$\frac{A}{a} \geq \frac{A_n}{a_n} v \cdot \bar{v}_n$$

for all n. Hence

$$\frac{A}{a} \geq \lim_{n\to\infty} \frac{A_n}{a_n} v \cdot \bar{v}_n = \frac{A^*(v)}{a^*(v)}.$$

But this implies $A \geq A^*(v)$ or $w \in W(A^*)$, a contradiction.

Hence $\lim_{n\to\infty} \mu(W_\varepsilon) = \mu(W^*)$. Therefore,

$$\lim_{n\to\infty} \frac{V(x_\varepsilon)}{\varphi\varepsilon_n} = \lim_{n\to\infty} \mu(W_\varepsilon) = \mu(W^*).$$

Taking the sequence referred to in the statement of the lemma to be $\{\varphi\varepsilon_n\}$ gives the desired conclusion.

Part 2. Now we drop the assumption that A^* is bounded. Fix a sequence $\{\varepsilon_n\}$ with $\varepsilon_n > 0$ for all n such that $\varepsilon_n \to 0$. Define a new function

$$A_n^*(v) = \begin{cases} A^*(v), & \text{if } -1+\varepsilon_n \leq A^*(v) \leq 1-\varepsilon_n \\ -1+\varepsilon_n, & \text{if } A^*(v) < -1+\varepsilon_n \\ 1-\varepsilon_n, & \text{if } A^*(v) > 1-\varepsilon_n \end{cases}$$

Clearly, A_n^* is bounded for every n. It is tedious but not difficult to show that the fact that $W(A^*)$ is closed implies that $W(A_n^*)$ is closed for every n. Hence for each n, we can find sequences $\{\varepsilon_m^n\}$ and $\{x_m^n\}$ such that

$$\lim_{m\to\infty} \frac{V(x_m^n)}{\varepsilon_m^n} = \mu(W(A_n^*)).$$

That is, for any $\delta > 0$, there exists $M_n(\delta)$ such that

$$\left| \frac{V(x_m^n)}{\varepsilon_m^n} - \mu(W(A_n^*)) \right| < \delta, \quad \forall m \geq M_n(\delta).$$

Rewriting,

$$\left| \frac{V(x_m^n)}{\varepsilon_m^n} - \mu(W(A^*)) + \mu(W(A^*) \setminus W(A_n^*)) - \mu(W(A_n^*) \setminus W(A^*)) \right| < \delta, \quad \forall m \geq M_n(\delta).$$

Clearly, though, if $w \in W(A^*) \setminus W(A_n^*)$, then there must be some N such that $w \in W(A^*) \cap W(A_{\bar{n}}^*)$ for all $\bar{n} \geq N$. The analogous statement is true for $W(A_n^*) \setminus W(A^*)$.

So fix any sequence $\{\delta_n\}$ converging to zero. For each n, fix $m_n \geq M_n(\delta_n)$. Consider the sequence $\{\hat{\varepsilon}_n\} = \{\varepsilon_{m_n}^n\}$ and $\{\hat{x}_n\} = \{x_{m_n}^n\}$. Clearly, for every n,

$$\left| \frac{V(\hat{x}_n)}{\hat{\varepsilon}_n} - \mu(W(A^*)) + k\mu(W(A^*) \setminus W(A_n^*)) - \mu(W(A_n^*) \setminus W(A^*)) \right| < \delta_n.$$

So

$$\lim_{n \to \infty} \frac{V(\hat{x}_n)}{\hat{\varepsilon}_n} = \mu(W(A^*)). \blacksquare$$

To complete the proof of Theorem 3, suppose (u, μ) and $(\bar{u}, \bar{\mu})$ are random Strotz representations of $\succeq$ where μ and $\bar{\mu}$ are measures over $\mathcal{W}$. Let V and $\bar{V}$ denote the utility functions over menus generated by (u, μ) and $(\bar{u}, \bar{\mu})$ respectively. As explained in the text, u and $\bar{u}$ must be the same up to a positive affine transformation. For convenience, we transform so that $u = \bar{u}$.

Since V and $\bar{V}$ are random Strotz representations of the same preference, there exists $a > 0$ and b such that $V(x) = a\bar{V}(x) + b$ for all menus x. For $x = \{\alpha\}$, then, $V(\{\alpha\}) = a\bar{V}(\{\alpha\}) + b$ or $u(\alpha) = a\bar{u}(\alpha) + b$. Since $u = \bar{u}$, then, $a = 1$ and $b = 0$. In other words, we must have $V = \bar{V}$.

Note that the sequence of menus constructed in the proof of Lemma 5 is independent of the representation. Hence for any set W which is closed and closed under C_u, we must have $\mu(W) = \bar{\mu}(W)$.

Fix any measurable set $E \subseteq \mathcal{W}$ which is closed under C_u. By Theorem 12.3 of Billingsley (1995, page 174),

$$\mu(E) = \sup_{W \subseteq E | W \text{ closed}} \mu(W)$$

and similarly for $\bar{\mu}$. It is easy to see that if $W \subseteq E$, then if we close W under C_u, the resulting set will be contained in E. That is,

$$W^* \equiv \{w \in \mathcal{W} \mid w \; C_u \; w', \text{ for some } w' \in W\} \subseteq E.$$

Obviously, $\mu(W^*) \geq \mu(W)$. Hence

$$\mu(E) = \sup_{W \subseteq E | W \text{ closed, closed under } C_u} \mu(W).$$

Since $\mu(W) = \bar{\mu}(W)$ for any W which is closed and closed under C_u, this implies $\mu(E) = \bar{\mu}(E)$. Thus μ and $\bar{\mu}$ coincide for any measurable set which is closed under C_u.

Let $\mathcal{P}$ denote the collection of measurable sets E which are closed under C_u. So we have established that μ and $\bar{\mu}$ coincide on $\mathcal{P}$. It is easy to show that $\mathcal{P}$ is a π–system. To see this, suppose $E_1, E_2 \in \mathcal{P}$. Then E_1 and E_2 are measurable, so $E_1 \cap E_2$ is measurable. Also, fix any $w \in E_1 \cap E_2$ and any w' such that $w' \; C_u \; w$. Since $w \in E_i$ and E_i is closed under C_u, we must have $w' \in E_i$, $i = 1, 2$. Hence $w' \in E_1 \cap E_2$. So $E_1 \cap E_2$ is closed under C_u and hence is an element of $\mathcal{P}$. Hence $\mathcal{P}$ is a π–system.

Theorem 3.3 of Billingsley (1995) then implies $\mu = \bar{\mu}$ on $\sigma(\mathcal{P})$, the σ–algebra generated by $\mathcal{P}$. We now show that $\sigma(\mathcal{P})$ is the Borel σ–algebra, completing the proof of uniqueness.

Fix any open set $W \subseteq \mathcal{W}$ and any $w \in W \setminus \{-u, u\}$. It is not hard to see that there exists a closed $\hat{\mathcal{V}} \subseteq \mathcal{V}$ and rational numbers $r_1, r_2 \in (-1, 1]$ such that $w \in W(A_1^*) \setminus W(A_2^*) \subseteq W$ where

$$A_i^*(v) = \begin{cases} r_i, & \text{for } v \in \hat{\mathcal{V}} \\ 1, & \text{otherwise} \end{cases}$$

Obviously, $W(A_1^*), W(A_2^*) \in \mathcal{P}$ implies $W(A_1^*) \setminus W(A_2^*) \in \sigma(\mathcal{P})$. Note that $\{u\} \in \mathcal{P}$ and that $\mathcal{W} \setminus \{-u\} \in \mathcal{P}$ implies $\{-u\} \in \sigma(\mathcal{P})$. Hence W is the union of a countable collection of sets in $\sigma(\mathcal{P})$ so $W \in \sigma(\mathcal{P})$. Hence $\sigma(\mathcal{P})$ contains all open sets and so contains the Borel σ–algebra. Since all sets in $\mathcal{P}$ are in the Borel σ–algebra, $\sigma(\mathcal{P})$ cannot be larger than the Borel σ–algebra, so it must equal it. ▌

C Proof of Theorem 4

Lemma 6. *Suppose $\succeq_i$ has a random Strotz representation (u, μ_i), $i = 1, 2$. Then $\succeq_2$ is more temptation averse than $\succeq_1$ if and only if for every menu x,*

$$V_2(x) \equiv \int \max_{\beta \in B_w(x)} u(\beta) \, \mu_2(dw) \leq \int \max_{\beta \in B_w(x)} u(\beta) \, \mu_1(dw) \equiv V_1(x).$$

Proof of Lemma. Suppose $\succeq_2$ is more temptation averse than $\succeq_1$ but that $V_2(x) > V_1(x)$ for some menu x. Without loss of generality, assume x is closed and convex. Then there exists $\alpha \in x$ such that $\{\alpha\} \sim_2 x$. So $u(\alpha) = V_2(x) > V_1(x)$. Hence $\{\alpha\} \succ_1 x$ but we do not have $\{\alpha\} \succ_2 x$, contradicting $\succeq_2$ more temptation averse. For the converse, suppose $V_1(x) \geq V_2(x)$ for all x. Then whenever $u(\alpha) > V_1(x)$, we have $u(\alpha) > V_2(x)$, so $\succeq_2$ is more temptation averse than $\succeq_1$. ▌

First, we show that if μ_2 conditionally–dominates μ_1, then $\succeq_2$ is more temptation averse than $\succeq_1$. Fix any menu and any $v \in \mathcal{V}$. Since the utility u gets from the choice is weakly increasing in A, we know that the expected utility of the menu conditional on v

is higher under μ_A^1 than under μ_A^2 for (almost) every v. Hence the same is true when we take expectations over v since the marginals are the same. Thus $V_1(x) \geq V_2(x)$ for all x, implying $\succeq_2$ is more temptation averse than $\succeq_1$ by Lemma 6.

For the converse, suppose $\succeq_2$ is more temptation–averse than $\succeq_1$. We construct decompositions of μ_1 and μ_2 which show that μ_2 conditionally–dominates μ_1. We define $\mathcal{V}$ as in the proof of Theorem 3. To begin constructing the relevant measures, we fix a partition $\mathcal{V}_1, \ldots, \mathcal{V}_N$ of $\mathcal{V}$ with the property that each $\mathcal{V}_i$ is measurable. We refer to such a partition as a *measurable partition*. For any $\mathcal{V}_n$ and any $A_n \in [-1, 1]$, let

$$\mu_i(A_n, \mathcal{V}_n) = \mu_i\left(\{w \in \mathcal{W} \mid w = v\sqrt{1-A^2} + Au, \text{ for } 1 > A \geq A_n,\ A \neq -1,\ v \in \mathcal{V}_n\}\right).$$

Note that $\mu_i(A_n, \mathcal{V}_n)$ is defined so that it does not include the measure of u or $-u$. In particular, $\mu_i(1, \mathcal{V}_n) = 0$ and

$$\mu_i(-1, \mathcal{V}_n) = \mu_i\left(L(\mathcal{V}_n)\right)) - \mu_i(\{u\}) - \mu_i(\{-u\})$$

where

$$L(\mathcal{V}_n) = \left\{w \in \mathcal{W} \mid w = v\sqrt{1-A^2} + Au, \text{ for } A \in [-1, 1],\ v \in \mathcal{V}_n\right\}.$$

Fix any $A_1, \ldots, A_N$ and define

$$W^* = \bigcup_{n=1}^{N} \bigcup_{v \in \mathcal{V}_\mathcal{N}} \{w \in \mathcal{W} \mid w = v\sqrt{1-A^2} + Au,\ A \geq A_n,\ A \neq -1\}.$$

First, suppose W^* is closed. By Lemma 5, we know that there is a sequence of menus x_m and numbers ε_m such that

$$\lim_{m \to \infty} \frac{V_i(x_m)}{\varepsilon_m} = \mu_i(W^*).$$

Since this sequence is independent of the preference (given that both have commitment utility u) and since $V_1(x_m) \geq V_2(x_m)$ for all m, we have $\mu_1(W^*) \geq \mu_2(W^*)$.

Now suppose that W^* is not closed. In this case, Theorem 12.3 of Billingsley (1995, page 174) implies that

$$\mu_2(W^*) = \sup_{W \subseteq E | W \text{ closed}} \mu_2(W).$$

As shown in the proof of Theorem 3, we can rewrite this as

$$\mu_2(W^*) = \sup_{W \subseteq W^* | W \text{ closed, closed under } C_u} \mu_2(W).$$

But we know that for every W which is closed and closed under C_u, we have $\mu_2(W) \leq \mu_1(W)$. Hence

$$\begin{aligned}\mu_2(W^*) &= \sup_{W \subseteq E | W \text{ closed}} \mu_2(W) \\ &\leq \sup_{W \subseteq W^* | W \text{ closed, closed under } C_u} \mu_1(W) \\ &= \mu_1(W^*).\end{aligned}$$

Summarizing, we have

$$\mu_1(u) + \sum_{n=1}^{N} \mu_1(A_n, \mathcal{V}_n) \geq \mu_2(u) + \sum_{n=1}^{N} \mu_2(A_n, \mathcal{V}_n). \tag{5}$$

for any $(A_1, \ldots, A_N) \in [-1, 1]^N$.

Note that this implies $\mu_1(u) \geq \mu_2(u)$ and $\mu_1(-u) \leq \mu_2(-u)$. The former is implied by taking $A_n = 1$ for all n and the latter by $A_n = -1$ for all n.

Let $\mu^*(u) = \mu_1(u) - \mu_2(u)$. First, assume $\mu^*(u) > 0$. For each n, let

$$\lambda_n^* = \sup_{A_n \in [-1,1]} \frac{\mu_2(A_n, \mathcal{V}_n) - \mu_1(A_n, \mathcal{V}_n)}{\mu^*(u)}.$$

By assumption, $\mu^*(u) > 0$, so this is well–defined. Also, note that for $A_n = 1$, the difference on the right–hand side is zero. Hence $\lambda_n^* \geq 0$.

Also,

$$\sum_n \lambda_n^* = \frac{1}{\mu^*(u)} \sum_n \sup_{A_n \in [-1,1]} [\mu_2(A_n, \mathcal{V}_n) - \mu_1(A_n, \mathcal{V}_n)].$$

Suppose this is strictly greater than 1. Then for each n, there is a sequence $\{A_n^m\}$ such that

$$\lim_{m \to \infty} \sum_n \mu_2(A_n^m, \mathcal{V}_n) > \mu^*(u) + \lim_{m \to \infty} \sum_n \mu_1(A_n^m, \mathcal{V}_n).$$

Substituting for $\mu^*(u)$ and rearranging,

$$\lim_{m \to \infty} \left[\mu_2(u) + \sum_n \mu_2(A_n^m, \mathcal{V}_n) \right] > \lim_{m \to \infty} \left[\mu_1(u) + \sum_n \mu_1(A_n^m, \mathcal{V}_n) \right].$$

But this contradicts equation (5). Hence $\sum_n \lambda_n^* \leq 1$.

Fix any $\lambda_1^1, \ldots, \lambda_N^1$, summing to 1, such that $\lambda_n^1 \geq \lambda_n^*$ for all n. Obviously, such a λ^1 exists. Then the definition of λ_n^* and the fact that $\lambda_n^1 \geq \lambda_n^*$ implies

$$\lambda_n^1 \mu^*(u) + \mu_1(A_n, \mathcal{V}_n) \geq \mu_2(A_n, \mathcal{V}_n), \quad \forall A_n \in [-1, 1], \ \forall n.$$

Substituting for $\mu^*(u)$, then,

$$\lambda_n^1 \mu_1(u) + \mu_1(A_n, \mathcal{V}_n) \geq \lambda_n^1 \mu_2(u) + \mu_2(A_n, \mathcal{V}_n), \quad \forall A_n \in [-1, 1], \ \forall n. \tag{6}$$

Next, suppose $\mu^*(u) = 0$. In this case, define $\lambda_n^1 = 1/N$ for $n = 1, \ldots, N$. Then equation (5) evaluated at any fixed A_n with $A_m = 1$ for all $m \neq n$ implies equation (6).

Next, define $\mu^*(-u) = \mu_2(-u) - \mu_1(-u)$. First, assume $\mu^*(-u) > 0$. Then define $\lambda_1^2, \ldots, \lambda_N^2$ by

$$\lambda_n^1 \mu_1(u) + \mu_1(-1, \mathcal{V}_n) = \lambda_n^1 \mu_2(u) + \mu_2(-1, \mathcal{V}_n) + \lambda_n^2 \mu^*(-u). \tag{7}$$

By equation (6) at $A_n = -1$, $\lambda_n^2 \geq 0$ for all n. Also, summing both sides over n and using $\sum_n \lambda_n^1 = 1$, we obtain

$$\mu_1(u) + \sum_n \mu_1(-1, \mathcal{V}_n) = \mu_2(u) + \sum_n \mu_2(-1, \mathcal{V}_n) + \mu^*(-u) \sum_n \lambda_n^2.$$

The left–hand side is $\mu_1(\mathcal{W}) - \mu_1(-u) = 1 - \mu_1(-u)$. The right–hand side is $1 - \mu_2(-u) + \mu^*(-u) \sum_n \lambda_n^2$. Hence we have

$$\mu^*(-u) = \mu^*(-u) \sum_n \lambda_n^2.$$

Since $\mu^*(-u) > 0$ by assumption, we must have $\sum_n \lambda_n^2 = 1$.

Second, suppose $\mu^*(-u) = 0$. In this case, the definition of λ^1 implies

$$\lambda_n^1 \mu_1(u) + \mu_1(-1, \mathcal{V}_n) \geq \lambda_n^1 \mu_2(u) + \mu_2(-1, \mathcal{V}_n) \tag{8}$$

for every n. Summing both sides over n and using $\sum_n \lambda_n^1 = 1$, we obtain

$$\mu_1(u) + \sum_n \mu_1(-1, \mathcal{V}_n) \geq \mu_2(u) + \sum_n \mu_2(-1, \mathcal{V}_n).$$

But since $\mu^*(-u) = 0$, we have $\mu_1(-u) = \mu_2(-u)$, so

$$\mu_1(u) + \sum_n \mu_1(-1, \mathcal{V}_n) + \mu_1(-u) \geq \mu_2(u) + \sum_n \mu_2(-1, \mathcal{V}_n) + \mu_2(-u).$$

But both each side of this inequality must equal 1. Hence equation (8) must hold with equality for all n. In light of this, we can define $\lambda_n^2 = 1/n$ for all n and equation (7) still holds.

This implies that we can rewrite μ_i as a measure $\hat{\mu}_i$ over $[-1,1] \times \mathcal{V}$, $i = 1, 2$, as follows. For any measurable $E \subseteq (-1,1) \times \mathcal{V}$, let

$$\hat{\mu}_i(E) = \mu_i\left(\left\{w \in \mathcal{W} \mid w = v\sqrt{1-A^2} + Au,\ (A,v) \in E\right\}\right).$$

For $E = \{1\} \times \mathcal{V}_n$, let

$$\hat{\mu}_i(E) = \lambda_n^1 \mu_i(u)$$

and for $E = \{-1\} \times \mathcal{V}_n$, let

$$\hat{\mu}_i(E) = \lambda_n^2 \mu_i(-u).$$

To see that such a measure exists, for each n, choose an arbitrary $v_n \in \mathcal{V}_n$ and assign probability $\lambda_n^1 \mu_i(u)$ to $\{1\} \times v_n$ and probability $\lambda_n^2 \mu_i(-u)$ to $\{-1\} \times v_n$. Extend this to the Borel σ–algebra on $[-1,1] \times \mathcal{V}$ in the obvious manner. That is, for each $E \subseteq [-1,1] \times \mathcal{V}_n$, let

$$\begin{aligned}\hat{\mu}_i(E) = \hat{\mu}_i(E \cap [(-1,1) \times \mathcal{V}]) + \hat{\mu}_i(E \cap (\{1\} \times \{v_i \mid i = 1, \ldots, N\})) \\ + \hat{\mu}_i(E \cap (\{-1\} \times \{v_i \mid i = 1, \ldots, N\})).\end{aligned}$$

The key point to observe about these measures is that for every n and every $A_n \in [-1,1]$, we have

$$\hat{\mu}_1([A_n, 1] \times \mathcal{V}_n) = \lambda_n^1 \mu_1(u) + \mu_1(A_n, \mathcal{V}_n) \geq \lambda_n^1 \mu_2(u) + \mu_2(A_n, \mathcal{V}_n) = \hat{\mu}_2([A_n, 1] \times \mathcal{V}_n)$$

and

$$\begin{aligned}\hat{\mu}_1([-1,1] \times \mathcal{V}_n) =& \lambda_n^1 \mu_1(u) + \mu_1(-1, \mathcal{V}_n) + \lambda_n^2 \mu_1(-u) \\ =& \lambda_n^1 \mu_2(u) + \mu_2(-1\mathcal{V}_n) + \lambda_n^2 \mu_2(-u) \\ =& \hat{\mu}_2([-1,1] \times \mathcal{V}_n).\end{aligned}$$

Generalizing, given any finite measurable partition Π of $\mathcal{V}$, let $\mathcal{M}_\Pi$ be the set of pairs of measures $(\hat{\mu}_1, \hat{\mu}_2)$ over $[-1,1] \times \mathcal{V}$ such that

$$\begin{aligned}\hat{\mu}_i(E) = \mu_i\left(\left\{w \in \mathcal{W} \mid w = v\sqrt{1-A^2} + Au,\ (A,v) \in E\right\}\right), \\ \forall \text{ measurable } E,\ i = 1, 2,\end{aligned} \tag{9}$$

$$\hat{\mu}_1([A_n, 1) \times \mathcal{V}_n) \geq \hat{\mu}_2([A_n, 1] \times \mathcal{V}_n), \quad \forall A_n \in [-1,1],\ \forall n, \tag{10}$$

and

$$\hat{\mu}_1([-1,1] \times \mathcal{V}_n) = \hat{\mu}_2([-1,1] \times \mathcal{V}_n), \quad \forall n. \tag{11}$$

We have shown that for every finite measurable partition Π, $\mathcal{M}_\Pi$ is nonempty. It is also not hard to see that it must be closed. Clearly, if Π' is a refinement of Π, then $\mathcal{M}_{\Pi'} \subseteq \mathcal{M}_\Pi$.

Each $\mathcal{M}_\Pi$ is a closed nonempty subset of the space of pairs of measures over $\mathcal{V}$, obviously a compact set. Fix a finite collection of finite measurable partitions, say $\Pi_1, \ldots, \Pi_T$. Let Π be the coarsest common refinement of these partitions. Then $\mathcal{M}_\Pi \subseteq \mathcal{M}_{\Pi_t}$ for all t. Since $\mathcal{M}_\Pi$ must be nonempty, we see that $\cap_t \mathcal{M}_{\Pi_t} \neq \emptyset$. By Kelly (1955, Chapter 5, Theorem 1), this implies that $\cap_\Pi \mathcal{M}_\Pi$ is nonempty where the intersection is taken over the set of all finite measurable partitions. Hence there is at least one pair of measures which satisfies equations (9), (10), and (11) for every finite measurable partition.

Hence we have shown that we can rewrite μ_1 and μ_2 as distributions $\hat{\mu}_1$ and $\hat{\mu}_2$ over $(A, v) \in [-1, 1] \times \mathcal{V}$ with the following properties. First, equation (9) implies that for every menu x,

$$\int_w \max_{\beta \in B_w(x)} u(\beta)\, \mu_i(dw) = \int_{(A,v)} \max_{\beta \in B_{v\sqrt{1-A^2}+Au}(x)} u(\beta)\, \hat{\mu}_i(d(A,v)), \quad i = 1, 2.$$

This holds since we have only specified how mass at u and $-u$ is spread across the sets $\{1\} \times \mathcal{V}$ and $\{-1\} \times \mathcal{V}$ respectively. Since $(1, v)$ and $(1, v')$ both correspond to utility function u, this has no effect on the calculation of the utility of any menu.

Second, equation (10) implies that for every measurable function $A^* : \mathcal{V} \to [-1, 1]$,

$$\int_v \hat{\mu}_1([A^*(v), 1] \times \{v\})\, dv \geq \int_v \hat{\mu}_2([A^*(v), 1] \times \{v\})\, dv.$$

To see this, simply note since A^* is bounded and measurable, there exists an increasing sequence of simple functions A^*_n converging to A^* pointwise from below.[21] Letting $W^* = \{(A, v) \mid A \geq A^*(v)\}$ and $W_n = \{(A, v) \mid A \geq A^*_n(v)\}$, we see that $W^* = \cap_n W_n$, so $\hat{\mu}_i(W^*) = \lim_{n\to\infty} \hat{\mu}_i(W_n)$. Hence

$$\hat{\mu}_1(W^*) = \lim_{n\to\infty} \hat{\mu}_1(W_n) \geq \lim_{n\to\infty} \hat{\mu}_2(W_n) = \hat{\mu}_2(W^*),$$

where the inequality follows from equation (10).

Third, it is easy to see that equation (11) implies that the marginals of $\hat{\mu}_1$ and $\hat{\mu}_2$ over $\mathcal{V}$ are the same.

[21] It is straightforward to modify the proof of Theorem 13.5 in Billingsley (1995, page 185) to show this.

Letting $\mu_{\mathcal{V}}$ denote the (common) marginal of $\hat{\mu}_1$ and $\hat{\mu}_2$ on $\mathcal{V}$ and $\mu_A^i(\cdot \mid v)$ a regular version of the conditional for $\hat{\mu}_i$, we see that $(\mathcal{V}, \mu_{\mathcal{V}}^i, \mu_A^i(\cdot \mid v))$ is a decomposition of μ_i for $i = 1, 2$.

Now we complete the proof by showing that $\mu_A^1(\cdot \mid v)$ FOSD $\mu_A^2(\cdot \mid v)$ for almost all v.

Let $\mathcal{V}(\bar{A})$ denote the set of v such that $\mu_A^1([\bar{A}, 1] \mid v) < \mu_A^2([\bar{A}, 1] \mid v)$ and let $\mathcal{V}^* = \cup_{\bar{A} \in [-1,1]} \mathcal{V}(\bar{A})$. We now show that $\mathcal{V}^*$ has $\mu_{\mathcal{V}}$ measure zero.

First, note that if there is an $\bar{A}$ such that $\mu_A^1([\bar{A}, 1] \mid v) < \mu_A^2([\bar{A}, 1] \mid v)$, then there must be a rational $\bar{A}$ with this property. This is obviously true if the distributions are continuous in a neighborhood of $\bar{A}$. If a distribution has a mass point at $\bar{A}$, then we can perturb the $\bar{A}$ slightly in one direction and maintain the inequality. Hence $\mathcal{V}^* = \cup_{\bar{A} \in \mathcal{R}} \mathcal{V}(\bar{A})$ where $\mathcal{R}$ denotes the rationals. For any $\bar{A}$, $\mathcal{V}(\bar{A})$ is measurable, so, as a countable union of measurable sets, $\mathcal{V}^*$ is measurable. Clearly,

$$\mu_{\mathcal{V}}(\mathcal{V}^*) \leq \sum_{\bar{A} \in \mathcal{R}} \mu_{\mathcal{V}} \left(\mathcal{V}(\bar{A}) \right).$$

To show that the right–hand side is zero, suppose it is positive. Then there must be some rational $\bar{A}$ such that $\mu_{\mathcal{V}}(\mathcal{V}(\bar{A})) > 0$. For every $v \in \mathcal{V}(\bar{A})$, we have

$$\mu_A^1([\bar{A}, 1] \mid v) \mu_{\mathcal{V}}(v) < \mu_A^2([\bar{A}, 1] \mid v) \mu_{\mathcal{V}}(v).$$

Integrating over $v \in \mathcal{V}(\bar{A})$, we get

$$\hat{\mu}_1([\bar{A}, 1] \times \mathcal{V}(\bar{A})) < \hat{\mu}_2([\bar{A}, 1] \times \mathcal{V}(\bar{A})),$$

a contradiction to equation (10). ∎

References

[1] Ahn, D., "Ambiguity without a State Space," *Review of Economic Studies*, **75**, January 2008, 3–28.

[2] Battaglini, M., R. Benabou, and J. Tirole, "Self–Control in Peer Groups," *Journal of Economic Theory*, **123**, August 2005, 105–134.

[3] Benabou, R., and M. Pycia, "Dynamic Inconsistency and Self-Control: A Planner–Doer Interpretation," *Economic Letters*, **77**, 2002, 419–424.

[4] Benabou, R., and J. Tirole, "Willpower and Personal Rules," *Journal of Political Economy*, **112**, August 2004, 848–887.

[5] Benabou, R., and J. Tirole, "Identity, Morals and Taboos: Beliefs as Assets," Princeton University working paper, June 2010.

[6] Billingsley, P., *Probability and Measure*, Third Edition, New York: Wiley, 1995.

[7] Caplin, A., and J. Leahy, "The Recursive Approach to Time Inconsistency," *Journal of Economic Theory*, **131**, November 2006, 134–156.

[8] Chatterjee, K., and R. V. Krishna, "Menu Choice, Environmental Cues and Temptation: A 'Dual Self' Approach to Self–Control," Pennsylvania State University working paper, 2007.

[9] Chatterjee, K., and R. V. Krishna, "A 'Dual Self' Approach to Stochastic Temptation," *American Economic Journal: Microeconomics*, **1**, August 2009, 148–167.

[10] Dekel, E., B. Lipman, and A. Rustichini, "Temptation–Driven Preferences," *Review of Economic Studies*, **76**, July 2009, 937–971.

[11] Dekel, E., and B. Lipman, "Costly Self Control and Random Self Indulgence," Boston University working paper, 2010.

[12] Eliaz, K., and R. Spiegler, "Contracting with Diversely Naive Agents," *Review of Economic Studies*, **73**, July 2006, 689–714.

[13] Epstein, L., "A Definition of Uncertainty Aversion," *Review of Economic Studies*, **66**, July 1999, 579–608.

[14] Fudenberg, D., and D. Levine, "A Dual Self Model of Impulse Control," *American Economic Review*, **96**, December 2006, 1449–1476.

[15] Fudenberg, D., and D. Levine, "Risk, Delay, and Convex Self–Control Costs," Harvard University working paper, 2010a.

[16] Fudenberg, D., and D. Levine, "Timing and Self–Control," Harvard University working paper, 2010b.

[17] Gajdos, T., T. Hayashi, J.-M. Tallon, and J.-C. Vergnaud, "Attitude toward Imprecise Information," *Journal of Economic Theory*, **140**, May 2008, 27–65.

[18] Ghirardato, P., and M. Marinacci, "Ambiguity Made Precise: A Comparative Foundation," *Journal of Economic Theory*, **102**, February 2002, 251–289.

[19] Gul, F., and W. Pesendorfer, "Temptation and Self–Control," *Econometrica*, **69**, November 2001, 1403–1435.

[20] Gul, F., and W. Pesendorfer, "Self–Control and the Theory of Consumption," *Econometrica*, **72**, January 2004, 119–158.

[21] Gul, F., and W. Pesendorfer, "Random Expected Utility," *Econometrica*, **74**, January 2006, 121–146.

[22] Harris, C., and D. Laibson, "Instantaneous Gratification," Harvard University working paper, June 2008.

[23] Higashi, Y., K. Hyogo, and N. Takeoka, "Subjective Random Discounting and Intertemporal Choice," *Journal of Economic Theory*, **144**, May 2009, 1015–1053.

[24] Kelly, J., *General Topology*, New York: Springer–Verlag, 1955.

[25] Laibson, D., "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, **112**, May 1997, 443–477.

[26] Lipman, B., and W. Pesendorfer, "Temptation," Boston University working paper, 2011.

[27] Mas–Colell, A., M. Whinston, and J. Green, *Microeconomic Theory*, New York: Oxford University Press, 1995.

[28] Milgrom, P., and I. Segal, "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, **70**, March 2002, 583–601.

[29] Noor, J., and N. Takeoka, "Uphill Self–Control," *Theoretical Economics*, **5**, May 2010a, 127–158.

[30] Noor, J., and N. Takeoka, "Menu–Dependent Self–Control," Boston University working paper, 2010b.

[31] O'Donoghue, Ted, and Matthew Rabin, "Doing It Now or Later," *American Economic Review*, **89**, March 1999, 103–124.

[32] Olszewki, W., "Preferences over Sets of Lotteries," *Review of Economic Studies*, **74**, April 2007, 567–595.

[33] Sarver, T., "Anticipating Regret: Why Fewer Options May Be Better," *Econometrica*, **76**, March 2008, 263–305.

[34] Stovall, J., "Multiple Temptations," *Econometrica*, **78**, January 2010, 349–376.

[35] Strotz, R., "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, **23**, 1955, 165–180.