

# Self-Control and the Theory of Consumption<sup>1</sup>

Faruk Gul

and

Wolfgang Pesendorfer

## Abstract

To study the behavior of agents who are susceptible to temptation in infinite horizon consumption problems under uncertainty, we define and characterize *dynamic self-control* (DSC) preferences. DSC preferences are recursive and separable. In economies with DSC agents, equilibria exist but may be inefficient; in such equilibria, steady state consumption is independent of initial endowments and increases in self-control. Increasing the preference for commitment while keeping self-control constant increases the equity premium. Removing non-binding constraints changes equilibrium allocations and prices. Debt contracts can be sustained even if the only feasible punishment for default is the termination of the contract.

KEYWORDS: Self-Control, preference for commitment, consumption behavior, time-inconsistency, borrowing constraints, risk-premium, sovereign debt.

# 1. Introduction

In experiments, subjects exhibit a reversal of preferences when choosing between a smaller, earlier and a larger, later reward (Kirby and Herrnstein (1995)). The earlier reward is preferred when it offers an immediate payoff whereas the later reward is preferred when both rewards are received with delay. This phenomenon is referred to as *time inconsistency* and has inspired theoretical work that modifies exponential discounting to allow for disproportionate discounting of the immediate future.<sup>2</sup>

This paper proposes an alternative approach to incorporate the experimental evidence. We extend our earlier analysis of self-control in two period decision problems (Gul and Pe-sendorfer (2001)) to an infinite horizon. Our goal is to reconcile the experimental evidence with tractable, dynamically consistent preferences and to apply the resulting model to the analysis of problems in macroeconomics and finance.

As an illustration, consider a consumption-savings problem. A consumer faces a constant interest rate  $r$  and has wealth  $b$  at the beginning of period 1. Each period, the consumer must decide how much to consume of his remaining wealth. Let  $z(b)$  denote the corresponding decision problem and let  $c$  denote the current consumption choice. Our axioms imply that the consumer has preferences of the form:

$$W(z(b)) = \max_{c \in [0, b]} \{u(c) + v(c) + \delta W(z(b')) - v(b)\}$$

where  $u$  and  $v$  are von Neumann-Morgenstern utility functions and  $b' = (b - c)(1 + r)$  is the wealth in the next period. These preferences describe an individual who, every period, is tempted to consume his entire endowment. Were he to do so, the term  $v(c) - v(b)$  would drop out. When he consumes less than his endowment, he incurs the disutility of self-control  $v(c) - v(b)$ . The utility function  $v$  represents “temptation”, that is, the individual’s urge for current consumption. Optimal behavior trades-off the temptation to consume with the long-run self-interest of the individual, represented by  $u + \delta W$ . The main theoretical result of this paper (Theorem 1) is a representation theorem yielding the utility function  $W$  above.

The preferences developed in this paper depend on what the individual consumes and on what he *could have consumed*. In a simple consumption-savings problem with no

liquidity constraints, the maximal temptation  $v(b)$  is the temptation utility of consuming all current wealth. More generally, the maximal temptation depends on the set of possible consumption choices for that period. For example, if there are borrowing constraints the set of possible consumption choices will reflect these constraints.

To allow a direct dependence of preferences on opportunity sets, we study preferences over *decision problems*. Building on work by Kreps and Porteus (1978), Epstein and Zin (1989) and Brandenburger and Dekel (1993) we develop the appropriate framework to study infinite horizon decision problems. With the set of decision problems as our domain, we define *Dynamic Self-Control (DSC)* preferences and derive a utility function for those preferences. The formula above is a special case of this derived utility function.

In the literature on dynamic inconsistency non-recursive preferences of the form  $u(c_1) + \sum_{t=2}^{\infty} \beta \delta^{t-1} u(c_t)$  are specified for the initial time period (Phelps-Pollak (1968), Laibson (1997)). In addition, it is assumed that the preferences governing behavior in period  $t$  are given by  $u(c_t) + \sum_{t'=t+1}^{\infty} \beta \delta^{t'-t} u(c_{t'})$ . Therefore at time  $t > 1$ , the individual's period  $t$  preferences differ from his *conditional* preferences - the preference over continuation plans implied by his first period preferences and choices prior to period  $t$ . This preference change is taken as a primitive to derive the individual's desire for commitment.

In contrast, we *start* with preferences that may exhibit a desire for commitment. The description of a "consumption path" includes both the actual consumption and what the individual could have consumed in each period. On this extended space, preferences are recursive and the conditional preferences are the same as period  $t$  preferences. Our representation theorem shows that DSC preferences can be interpreted as describing an individual whose temptation utility,  $v$ , interferes with his long-run self-interest represented by  $u + \delta W$ . The consumer uses self-control to mediate between his temptation utility and his long-run self-interest.

The recursive structure of our model allows us to apply standard techniques of dynamic programming to find optimal solutions. At the same time, the model is consistent with the type of preference reversal documented in the experimental literature. Consider a consumer who must decide between a smaller reward in period 1 and a larger reward in period 2. Since the earlier reward would lead to a larger consumption in the decision period, the

consumer incurs a self-control cost when opting for the later reward. Now consider the situation where the earlier reward is for period  $t > 1$  and the later for period  $t + 1$ . Since the decision is taken in period 1, the choice does not affect consumption in the decision period and therefore the later reward can be taken without incurring a self-control cost. Thus, the agent's behavior appears non-recursive when we only observe his consumption choices.

Our framework is rich enough to accommodate the infinite horizon, stochastic dynamic programming problems central to macroeconomics and financial economics. Techniques developed in those areas can be applied to DSC preferences to explore how self-control and preference for commitment change the conclusions of standard (macro)economic models. Section 5 contains applications of DSC preferences to competitive economies. In section 5.1, we analyze a deterministic exchange economy with DSC preferences. We give conditions under which a competitive equilibrium exists and find that in general, competitive equilibria are not Pareto efficient. We also examine steady state equilibria and show that with DSC preferences the steady state distribution of wealth is independent of the initial wealth distribution. In section 5.2, we observe that removing constraints that are not binding given the original equilibrium allocations and prices, may change these allocations and prices. In section 5.3, we analyze a simple stochastic exchange economy and find that (under appropriate assumptions on  $u$  and  $v$ ) increasing the preference for commitment increases the predicted premium of risky over safe assets. Section 6 explores incentive compatible debt contracts. If the only punishment for default is exclusion from future borrowing, then standard preferences imply that there are no incentive compatible debt contracts (Bulow and Rogoff (1989)). In contrast, with DSC preferences, "natural" debt contracts turn out to be incentive compatible.

## 2. Stationary Self-Control Preferences

Consider a decision-maker (DM) who must take an action in every period  $t = 1, 2, \dots$ . Each action results in a consumption for that period and constrains future actions. The standard approach to this problem is to define preferences for the DM over sequences of consumption realizations. This standard approach excludes preferences that depend not only on outcomes but also on what could have been chosen. Such a direct dependence on the opportunity set is natural in a context where individuals suffer from self-control problems.

To allow for preferences to depend on opportunity sets we make *decision problems* the domain of our preferences. Let  $C$  denote the compact metric space of possible consumptions in each period. An infinite-horizon consumption problem (IHCP) is a set of choices, each of which yields a consumption  $c \in C$  for the current period and an infinite horizon problem starting next period. Choices may yield a random consumption and continuation problem. Hence, an IHCP is a set of probability measures where each realization yields a consumption today and a continuation IHCP.

Let  $Z$  be the set of IHCPs. Each  $z \in Z$  can be identified with a compact set of probability measures on  $C \times Z$ . For any compact metric space  $X$ , let  $\Delta(X)$  denote the set of probability measures on  $X$  and let  $\mathcal{K}(X)$  denote the set of non-empty compact subsets of  $X$ . We endow  $\Delta(X)$  with the weak topology and note that  $\Delta(X)$  is compact and metrizable.<sup>3</sup> We endow  $\mathcal{K}(X)$  with the Hausdorff topology which implies that  $\mathcal{K}(X)$  is a compact metric space.<sup>4</sup> The domain of preferences  $Z$  can be identified with  $\mathcal{K}(\Delta(C \times Z))$ . Appendix A1 shows that  $Z$  is well-defined and a compact metric space.

We use  $x, y$  or  $z$  to denote elements of  $Z$ . When there is no risk of confusion, we write  $\Delta$  instead of  $\Delta(C \times Z)$ . We use  $\mu, \nu$  or  $\eta$  to denote elements of  $\Delta$ . A lottery that yields the current consumption  $c$  and the continuation problem  $z$  with certainty is denoted  $(c, z)$ .

For  $\alpha \in [0, 1]$ , let  $\alpha\mu + (1 - \alpha)\nu \in \Delta$  be the measure that assigns  $\alpha\mu(A) + (1 - \alpha)\nu(A)$  to each  $A$  in the Borel  $\sigma$ -algebra of  $C \times Z$ . Similarly,  $\alpha x + (1 - \alpha)y := \{\alpha\mu + (1 - \alpha)\nu \mid \mu \in x, \nu \in y\}$  for  $\alpha \in [0, 1]$  denotes the convex combination of  $x$  and  $y$ .

A preference is denoted  $\succeq$ . We make the following familiar assumptions on preferences.

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

**Axiom 2:** (Continuity) The sets  $\{x \mid x \succeq z\}$  and  $\{x \mid z \succeq x\}$  are closed.

We say that the function  $W : Z \rightarrow \mathbb{R}$  represents the preference  $\succeq$  when  $x \succeq y$  iff  $W(x) \geq W(y)$ . Axioms 1 and 2 imply that  $\succeq$  may be represented by a continuous function  $W$ .

A standard DM evaluates a set of options by its best element. Adding options to the set can never make such a DM worse off. Thus, if  $x \succeq y$  then the best option in  $x$  is preferred to the best option in  $y$  and hence

$$x \succeq y \Rightarrow x \cup y \sim x \tag{*}$$

Axioms 1, 2 together with (\*) imply that there is continuous utility function  $U$  such that  $W(x) := \max_{\mu \in x} U(\mu)$  represents  $\succeq$ .<sup>5</sup> Hence, Axioms 1, 2 and (\*) yield a standard DM. In contrast, a DM who is susceptible to temptation may prefer a smaller set of options to a larger set. That is, he may have a preference for commitment.

**Definition:** The preference  $\succeq$  has a preference for commitment at  $z$  if there is  $x \subset z$  such that  $x \succ z$ .

When  $\{\mu\} \succ \{\mu, \nu\} \sim \{\nu\}$  the DM is worse off when  $\nu$  is available and, in addition, derives no benefit from the availability of  $\mu$ . We interpret this as a situation where the DM succumbs to the temptation presented by  $\nu$ .

When  $\{\mu\} \succ \{\mu, \nu\} \succ \{\nu\}$  the DM derives benefit from the availability of  $\mu$  in spite of the presence of  $\nu$ . We interpret this as a situation where  $\nu$  presents a temptation but the DM exercises costly self-control and chooses  $\mu$  in the presence of  $\nu$ . More generally, we interpret  $x \succ x \cup y \succ y$  as a situation where the DM has self-control. Theorem 3 (Section 3) shows that, within the context of preferences analyzed here, this definition agrees with the everyday meaning of the term “self-control” as the ability to resist temptation.

**Definition:** The preference  $\succeq$  has self-control at  $z$  if there are subsets  $x, y$  with  $x \cup y = z$  and  $x \succ z \succ y$ .

Axiom 3 allows both a preference for commitment and self-control.

**Axiom 3:** (Set Betweenness)  $x \succeq y$  implies  $x \succeq x \cup y \succeq y$ .

Our next objective is to characterize preferences that have self-control and a stationary, separable representation. A singleton set represents a situations where the DM has no choice. Hence, we refer to the restriction of  $\succeq$  to singleton sets as the commitment ranking. The following axiom ensures that the commitment ranking satisfies von Neumann-Morgenstern's independence axiom.

**Axiom 4:** (*Independence*)  $\{\mu\} \succ \{\nu\}$ ,  $\alpha \in (0, 1)$  implies  $\{\alpha\mu + (1-\alpha)\eta\} \succ \{\alpha\nu + (1-\alpha)\eta\}$ .

Recall that  $(c, z)$  denotes a lottery that returns the consumption  $c$  in the current period and the continuation problem  $z$ . Axiom 5 requires that the correlation between the current consumption and the continuation problem does not affect preferences. The axiom considers two lotteries:  $\mu = \frac{1}{2}(c, z) + \frac{1}{2}(c', z')$  returns either the consumption  $c$  together with the continuation problem  $z$  or the consumption  $c'$  together with the continuation problem  $z'$ ;  $\nu = \frac{1}{2}(c, z') + \frac{1}{2}(c', z)$  returns either the consumption  $c$  together with the continuation problem  $z'$  or the consumption  $c'$  together with the continuation problem  $z$ . The axiom requires the DM to be indifferent between  $\{\mu\}$  and  $\{\nu\}$ .

**Axiom 5:** (*Separability*)  $\{\frac{1}{2}(c, z) + \frac{1}{2}(c', z')\} \sim \{\frac{1}{2}(c, z') + \frac{1}{2}(c', z)\}$ .

Axiom 6 requires preferences to be stationary. Consider the lotteries,  $(c, x), (c, y)$  each leading to the same consumption in the current period. The axiom requires that  $\{(c, x)\}$  is preferred to  $\{(c, y)\}$  if and only if the continuation problem  $x$  is preferred to the continuation problem  $y$ .

**Axiom 6:** (*Stationarity*)  $\{(c, x)\} \succeq \{(c, y)\}$  iff  $x \succeq y$ .

Axiom 7 requires the individual to be indifferent to the timing of resolution of uncertainty. In the standard model this indifference is implicit in the assumption that the domain of preferences is the set of lotteries over consumption paths. The richer domain used in this paper permits agents who are not indifferent to the timing of resolution of uncertainty.<sup>6</sup> Since our purpose is to focus on temptation and self-control, we rule out preference for early or late resolution of uncertainty.

To understand Axiom 7, consider the lotteries  $\mu = \alpha(c, x) + (1 - \alpha)(c, y)$  and  $\nu = (c, \alpha x + (1 - \alpha)y)$ . The lottery  $\mu$  returns the consumption  $c$  together with the continuation

problem  $x$  with probability  $\alpha$  and the consumption  $c$  with the continuation problem  $y$  with probability  $1 - \alpha$ . In contrast,  $\nu$  returns  $c$  together with the continuation problem  $\alpha x + (1 - \alpha)y$  with probability 1. Hence,  $\mu$  resolves the uncertainty about  $x$  and  $y$  in the current period whereas  $\nu$  resolves this uncertainty in the future. If  $\{\mu\} \sim \{\nu\}$  then the DM is indifferent as to the timing of the resolution of uncertainty.

**Axiom 7:** (*Indifference to Timing*)  $\{\alpha(c, x) + (1 - \alpha)(c, y)\} \sim \{(c, \alpha x + (1 - \alpha)y)\}$ .

Axiom 8 requires that two alternatives,  $\nu, \eta$ , offer the same temptation if they have the same marginal over current consumption. For any  $\mu \in \Delta(C \times Z)$ ,  $\mu^1$  denotes its marginal on the first coordinate (current consumption) and  $\mu^2$  its marginal on the second coordinate (the continuation problem).

**Axiom 8:** (*Temptation by Immediate Consumption*)  $\nu^1 = \eta^1$ ,  $\{\mu\} \succ \{\mu, \nu\} \succ \{\nu\}$  and  $\{\mu\} \succ \{\mu, \eta\} \succ \{\eta\}$  implies  $\{\mu, \nu\} \sim \{\mu, \eta\}$ .

To understand Axiom 8, recall that  $\{\mu\} \succ \{\mu, \nu\} \succ \{\nu\}$  represents a situation where the DM is tempted by  $\nu$  but uses self-control and chooses  $\mu$ . Similarly,  $\{\mu\} \succ \{\mu, \eta\} \succ \{\eta\}$  means that the DM is tempted by  $\eta$  but chooses  $\mu$ . Hence, both situations lead to the same choice. According to Axiom 8, if  $\nu^1 = \eta^1$  then the DM is indifferent between the two situations and hence experiences the same temptation in both situations.

We call a preference “degenerate” if it never benefits from additional options. Thus  $\succeq$  is degenerate if for every IHCP  $x$  and  $y$ ,  $x \succeq x \cup y$ .

**Definition:** The preference  $\succeq$  is non-degenerate if there exists  $x, y$  such that  $y \subset x$  and  $x \succ y$ .

Theorem 1 provides a recursive and separable representation of non-degenerate preferences that satisfy Axioms 1-8.

**Theorem 1:** If the non-degenerate preference  $\succeq$  satisfies Axioms 1–8 then there is some  $\delta \in (0, 1)$ , continuous functions  $u, v : C \rightarrow \mathbb{R}$  and a continuous function  $W$  that represents  $\succeq$  such that

$$W(z) := \max_{\mu \in z} \int (u(c) + v(c) + \delta W(z')) d\mu(c, z') - \max_{\nu \in z} \int v(c) d\nu(c, z')$$



for all  $z \in Z$ . Conversely, for any  $\delta \in (0, 1)$ , continuous  $u, v : C \rightarrow \mathbb{R}$ , there is a unique continuous function  $W$  that satisfies the equation above and the preference it represents satisfies Axioms 1 – 8.

**Proof:** See Appendix.

Note that if  $W$  satisfies the equation in Theorem 1 for some  $\delta, u, v$ , then the preference represented by  $W$  is continuous only if  $W$  is continuous. Therefore, by Theorem 1 above, there is a unique preference that satisfies Axioms 1 – 8 for any  $\delta \in (0, 1)$  and continuous  $u, v$ . Theorem 1 extends our earlier representation for two-period decision problems. To see the relationship to the earlier work, note that Axioms 4, 6 and 7 imply the following, stronger version of the independence axiom:<sup>7</sup>

**Axiom 4\*:**  $x \succ y, \alpha \in (0, 1)$  implies  $\alpha x + (1 - \alpha)z \succ \alpha y + (1 - \alpha)z$ .<sup>8</sup>

In Gul and Pesendorfer (2001) we show that Axioms 1-3 and 4\* imply that the preference can be represented by a utility function  $W$  of the form

$$W(z) = \max_{\mu \in z} \{U(\mu) + V(\mu)\} - \max_{\nu \in z} V(\nu)$$

where  $W, U$  and  $V$  are linear functions. The contribution of Theorem 1 is to establish that  $U$  and  $V$  are of the following form:

$$\begin{aligned} U(\mu) &= \int u(c) + \delta W(x) d\mu(c, z) \\ V(\mu) &= \int v(c) d\mu(c, z) \end{aligned} \tag{1}$$

To explain the role of the axioms, we outline the main steps in the proof of equation (1). Axiom 5 implies that  $U$  is separable, that is

$$U(\mu) = \int \left( u(c) + \tilde{W}(z') \right) d\mu(c, z')$$

for some  $u$  and  $\tilde{W}$ . By Axiom 6,  $\tilde{W}$  must represent the same preference as  $W$ , that is,  $\tilde{W}(x) \geq \tilde{W}(y)$  if and only if  $W(x) \geq W(y)$ . Axiom 7 implies that  $\tilde{W}$  is linear, i.e.,  $\tilde{W}(\alpha x + (1 - \alpha)y) = \alpha \tilde{W}(x) + (1 - \alpha) \tilde{W}(y)$ . Since  $W$  is linear, the uniqueness of von Neumann-Morgenstern utility functions implies that  $\tilde{W} = \beta + \delta W$  for some  $\beta, \delta \in \mathbb{R}, \delta > 0$ .

Let  $x, y$  be two IHCPs that offer commitment to the consumption  $c$  in the first  $T$  periods. The IHCP  $x$  yields the continuation problem  $x'$  in period  $T + 1$  whereas  $y$  yields the continuation problem  $y'$  in period  $T + 1$ . If  $W$  is not constant, we may choose  $x', y'$  so that  $W(x') \neq W(y')$ . Note that  $x \rightarrow y$  as  $T \rightarrow \infty$  and hence continuity requires that  $\delta^{T+1}(W(x') - W(y')) \rightarrow 0$  as  $T \rightarrow \infty$ . Clearly, this implies that  $\delta < 1$ . We conclude that  $U(\mu) = \int (u(c) + \delta W(y)) d\mu(c, y)$  for some  $u : C \rightarrow \mathbb{R}$  and  $\delta \in (0, 1)$ .

Axiom 8 considers a situation where  $W(\{\mu, \nu\}) = U(\mu) + V(\mu) - V(\nu)$  and  $W(\{\mu, \eta\}) = U(\mu) + V(\mu) - V(\eta)$  and requires that  $W(\{\mu, \nu\}) = W(\{\mu, \eta\})$  when  $\mu$  and  $\nu$  yield the same current consumption. But this implies that  $V(\nu) = V(\eta)$  when  $\nu$  and  $\eta$  yield the same current consumption and hence  $V(\nu) = \int v(c) d\nu(c, z)$  for some  $v : C \rightarrow \mathbb{R}$ . Non-degeneracy is used to establish the existence of the alternatives  $\mu, \nu, \eta$  with the desired properties. Note that Axiom 8 is only used to establish the particular form of the temptation utility. Without it and without the non-degeneracy assumption, we would get an analogous representation with an unrestricted linear temptation utility  $V : \Delta \rightarrow \mathbb{R}$ .

We refer to non-degenerate preferences that satisfy Axioms 1 – 8 as *dynamic self control* (DSC) preferences. If some  $W$  of the form given in Theorem 1 represents the preference relation  $\succeq$ , we refer to the corresponding  $(u, v, \delta)$  as a representation of  $\succeq$  and sometimes as *the* preference  $(u, v, \delta)$ .

The next result establishes that the representation provided in Theorem 1 is unique in the sense that  $(u, v, \delta)$  and  $(u', v', \delta')$  represent the same preference if and only  $\delta = \delta'$ , and  $u', v'$  are a common affine transformation of  $u, v$ , that is,

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \alpha \cdot \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} \beta_u \\ \beta_v \end{pmatrix}$$

for some  $\alpha > 0, \beta_u, \beta_v \in \mathbb{R}$ .

**Theorem 2:** *Let  $\succeq$  be a DSC preference with a preference for commitment at some  $z \in Z$  and let  $(u, v, \delta)$  be a representation of  $\succeq$ . Then,  $(u', v', \delta')$  also represents  $\succeq$  if and only if  $\delta = \delta'$  and there exist  $\alpha > 0, \beta_u, \beta_v \in \mathbb{R}$  such that  $u' = \alpha u + \beta_u$  and  $v' = \alpha v + \beta_v$ .*

Theorem 2 assumes that  $\succeq$  is a DSC preference with a preference for commitment at some  $z$ . This implies that if  $(u, v, \delta)$  represents  $\succeq$  then  $u$  and  $v$  are not constant.<sup>9</sup> (A DSC

preference is non-degenerate and hence  $u$  is not constant. Since  $\succeq$  has a preference for commitment at some  $z$ , it follows that  $v$  is not constant.) When  $v$  is constant, the DM has a “standard” preference that can be represented by maximization of a discounted sum of utilities. When  $u$  is constant, the DM is indifferent between all decision problems. In both cases, the representation is not unique in the sense of Theorem 2. When  $u$  is constant, any choice of  $v$  yields the same preference. When  $v$  is constant, replacing  $v$  with  $\alpha u + \beta v$  for any  $\alpha \geq 0$  yields the same (standard) preference. In particular, a constant  $v$  and  $v = u$  are indistinguishable cases.

### 3. Choice Behavior and Interpretation

The function  $W$  represents the agent’s preference over decision problems. It is analogous to the indirect utility function in static demand theory or the value function in a dynamic programming problem. Period 0 choices among IHCP’s maximize  $W$ . Hence, in period 0 the agent chooses  $y$  over  $x$  if

$$W(y) \geq W(x)$$

In period 1, the agent must choose an alternative  $\mu$  from a decision problem  $x$ . To simplify the notation below, let  $E_\mu[f]$  denote the expectation of the (Borel measurable) function  $f : C \times Z \rightarrow \mathbb{R}$  with respect to  $\mu$ , i.e.,  $E_\mu[f] := \int f(c, z)d\mu(c, z)$ . We assume that the agent’s utility from choosing  $\mu \in x$  is

$$U(\mu) + V(\mu) - \max_{\eta \in x} V(\eta) = E_\mu [u + v + \delta W] - \max_{\eta \in x} E_\eta [v] \quad (2)$$

Note that the term  $\max_{\eta \in x} E_\eta [v]$  is independent of the choice and therefore the agent chooses  $\mu \in x$  to solve

$$\max_{\mu \in x} E_\mu [u + v + \delta W]$$

The indicated behavior for periods  $t > 0$  is an assumption because the representation asserts only that the DM assigns utility to decision problems *as if* he were behaving in this manner. The difficulty stems from the fact that the domain of preferences is *decision problems* and not choices from those problems. This difficulty can be overcome by

extending preferences to choices from decision problems. In Gul and Pesendorfer (2001) we provide such an extension and give conditions that ensure that the DM indeed behaves as suggested by the representation. Here, we simply assume that the DM behaves in this way.

Expression (2) extends the agent's utility function to period  $t \geq 1$  choices in a dynamically consistent manner. This means that from the perspective of any period  $t$ , choices in periods  $t' < t$  were made optimally and choices in periods  $t' > t$  will be made optimally. For example, assume that  $W(\{\mu\}) > W(\{\mu, \nu\})$  but  $E_\nu[u + v + \delta W] > E_\mu[u + v + \delta W]$ . This implies that the agent would have been better-off if alternative  $\nu$  were to be eliminated from his choice set but nevertheless chooses  $\nu$  from  $\{\mu, \nu\}$  in period 1. This apparent inconsistency is reconciled by allowing the utility to depend on the decision problem. The utility of choosing  $\mu$  from  $\{\mu\}$  is

$$E_\mu[u + v + \delta W] - \max_{\eta \in \{\mu\}} E_\eta[v] = E_\mu[u + \delta W]$$

When  $E_\nu[v] > E_\mu[v]$  the utility of choosing  $\mu$  from  $\{\mu, \nu\}$  is

$$E_\mu[u + v + \delta W] - \max_{\eta \in \{\mu, \nu\}} E_\eta[v] = E_\mu[u + \delta W] - (E_\nu[v] - E_\mu[v])$$

and the utility of choosing  $\nu$  from  $\{\mu, \nu\}$  is

$$E_\nu[u + v + \delta W] - \max_{\eta \in \{\mu, \nu\}} E_\eta[v] = E_\nu[u + \delta W]$$

We interpret the direct dependence of welfare on the decision problem as capturing an agent who struggles with temptations. An agent facing the IHCP  $\{\mu\}$  is committed to choosing  $\mu$  and hence experiences no temptation. When the agent chooses  $\mu$  from  $\{\mu, \nu\}$  he must exercise costly self-control because  $\nu$  presents a temptation. The cost of self-control is given by  $E_\nu[v] - E_\mu[v]$ . When this cost is large enough the optimal behavior is to succumb to the temptation presented by  $\nu$ .

An alternative interpretation would be to assert a *change in the agent's preference* between periods. According to this interpretation, each period is occupied by a distinct decision-maker with an independent utility function governing choice. The period 0 utility

of decision problem  $x$  is  $W(x)$  and period  $t \geq 1$  utility of  $\mu \in x$  is  $E_\mu[u + v + \delta W]$ . According to this interpretation, welfare in period  $t \geq 1$  depends only on the choice  $\mu$  and is otherwise independent of the decision problem.

We adopt the dynamically consistent interpretation for the following reason. The period 0 preference suggests that the agent wishes to avoid temptations even if they do not alter the choice. For example, if  $E_\mu[u + v + \delta W] > E_\nu[u + v + \delta W]$  and  $E_\nu[v] > E_\mu[v]$  the agent would be better off if  $\nu$  were eliminated from the set  $\{\mu, \nu\}$  even though he chooses  $\mu$  from  $\{\mu, \nu\}$  in period 1. We see no reason why this dependence of welfare on the decision problem should suddenly disappear as we change periods. The changing tastes interpretation assumes a choice-problem independence of welfare that seems to contradict the very premise of a model of self-control.

To use an analogy from a more familiar setting, consider a decision-maker who must choose between apple and banana in periods 0 and 1. The preference is

$$(\text{apple, banana}) \succ (\text{banana, banana}) \succ (\text{banana, apple}) \succ (\text{apple, apple})$$

Hence, in period 1 the agent always chooses banana. How would the agent rank (apple, banana) and (banana, banana) in period 1? Advocates of the ‘changing tastes’ view could argue that because in period 1 the agent always chooses the banana, he has no taste for variety in period 1. As a result, (banana, banana) is preferred to (apple, banana) in period 1. By contrast, the standard interpretation asserts the same ranking of pairs for both periods. After all, the period 0 ranking of pairs suggests that the agent has a preference for variety as long as banana is chosen in period 1 and period 1 behavior offers no evidence contradicting this preference. We feel similarly about preference for commitment. If in period 0 the agent chooses  $y \subset x$  over  $x$  then the period 1 choice behavior provides no evidence that the agent would not prefer commitment to  $y$  also in period 1. Hence, we see no reason to assert a change in preference.<sup>10</sup>

Our model assumes that period 0 is “special” in the sense that it is a period *prior* to the experience of temptation. Choices in all other periods are potentially affected by temptations. This suggests that correct identification of the parameters of our model requires a correct specification of period 0. However, immediate temptation together with

stationarity allows us to circumvent this issue and treat any period as “period 0” as long as we hold consumption in that period fixed. Assume that the agent faces the decision problem  $x = \{\mu, \nu\}$  where  $\mu = (c, y), \nu = (c, z)$ . Hence, the agent must choose between two alternatives that both yield the consumption  $c$  in period 1 but differ in the decision problem starting in period 2.<sup>11</sup> The agent chooses  $(c, x)$  from  $z$  if

$$u(c) + v(c) + \delta W(x) \geq u(c) + v(c) + \delta W(y)$$

which is equivalent to  $W(x) \geq W(y)$ . Since  $W$  represents the period 0 preference it follows that the period 1 choice among alternatives of the form  $(c, x), (c, y)$  reveals the period 0 preference between  $x$  and  $y$ . Therefore, we can interpret period 0 simply as *any* period in which consumption is unaffected by the individual’s choice.

When a decision problem offers commitment to a particular choice  $\mu$ , the agent’s utility is  $W(\{\mu\}) = E_\mu[u + \delta W]$ . Therefore, we refer to  $u + \delta W$  as the commitment utility of  $\mu$ . We refer to  $E_\mu[v]$  as the temptation utility of  $\mu$ . The DM prefers  $y \subset x$  to  $x$  only if  $\max_{\mu \in x} E_\mu[v] > \max_{\mu \in y} E_\mu[v]$ . Hence, commitment to  $y$  is desirable because  $x$  contains alternatives that are more tempting than the most tempting element of  $y$ .

Section 2 provided a definition of self-control based on the period 0 preference over decision problems. Here we provide an alternative definition based on choice behavior in period 1. If the agent chooses an alternative  $\mu$  from  $x = \{\mu, \nu\}$  and  $E_\mu[v] < E_\nu[v]$  it means that he did not choose the most tempting alternative. In other words, the agent exercised self-control. For any (Borel measurable) function  $f : C \times Z \rightarrow \mathbb{R}$ , let  $\mathcal{C}(z, f) := \{\mu \in z \mid E_\mu[f] \geq E_\nu[f] \text{ for all } \nu \in z\}$ . Hence, the set  $\mathcal{C}(z, u + v + \delta W)$  denotes the choices from  $z$  whereas  $\mathcal{C}(z, v)$  denotes the set of most tempting alternatives in  $z$ . The agent exercises self-control if  $\mathcal{C}(z, v) \cap \mathcal{C}(z, u + v + \delta W) = \emptyset$ . Theorem 3 below shows that for DSC preferences our definition of self-control based on period 0 preference (given in section 2) is equivalent to the definition of self-control based on behavior in period  $t = 1$ .

**Theorem 3:** *(Gul and Pesendorfer (2001)) The DSC preference  $(u, v, \delta)$  has self-control at  $z$  if and only if  $\mathcal{C}(z, v) \cap \mathcal{C}(z, u + v + \delta W) = \emptyset$ .*

**Proof:** See Theorem 2 of Gul and Pesendorfer 2001.

## 4. Preference Reversal

Experimental evidence on time preference has been the main motivation for research on preference for commitment. In a typical experiment a subject is asked to choose between a smaller, period  $t$  reward and a larger, period  $t + 1$  reward. Subjects tend to resolve this trade-off differently depending on *when* the decision is made. In particular, if the decision is made in period  $t$  then the smaller period  $t$  reward is chosen whereas the larger period  $t + 1$  reward tends to be chosen when the decision is made in period  $t - 1$  or earlier. (See, for example, Kirby and Herrnstein (1995)). The example below demonstrates that such a reversal is consistent with DSC preferences.

Suppose there is one good and  $C = [0, 1]$ . Further assume that  $u$  and  $v$  are increasing. To translate the experimental setting into our notation we describe decision problems in which the agent makes exactly one choice between a period  $t$  reward  $\alpha$  and a period  $t + 1$  reward  $\beta$  and is otherwise committed to a fixed consumption. First, consider the case where this decision is made in period  $t$ . Let  $z_0$  denote the IHCP in which the only option is to consume  $c$  each period and let  $z_\beta := \{(c + \beta, z_0)\}$  denote the IHCP in which the agent is committed to a “reward”  $\beta > 0$  in the initial period and the consumption  $c$  thereafter. The decision problem

$$z := \{(c, z_\beta), (c + \alpha, z_0)\}$$

with  $\beta > \alpha > 0$  describes the period  $t$  choice between a reward  $\alpha$  in period  $t$  and a reward  $\beta > \alpha$  in period  $t + 1$ .

Next, consider the case where the decision between the period  $t$  reward  $\alpha$  and the period  $t + 1$  reward  $\beta$  is made in period  $t - 1$ . In that case, the decision problem is

$$z' := \{(c, \{(c, z_\beta)\}), (c, \{(c + \alpha, z_0)\})\}$$

Note that both choices from  $z'$  lead to the consumption  $c$  in the decision period. The two alternatives are distinguished by the implied period  $t$  decision problems. The reward  $\alpha$  implies the period  $t$  decision problem  $\{(c + \alpha, z_0)\}$  whereas the reward  $\beta$  implies the period  $t$  decision problem  $\{(c, z_\beta)\}$ . Even though  $z$  and  $z'$  describe the same intertemporal trade-off they constitute different decision problems because the decision is made at a different date.

From the decision problem  $z$ , the agent chooses the earlier reward  $\alpha$  if

$$u(c + \alpha) + \delta W(z_0) > u(c) + v(c) - v(c + \alpha) + \delta W(z_\beta)$$

If the agent were to choose  $(c + \alpha, z_0)$  from  $z$  he would incur no self-control cost whereas if the agent were to choose  $(c, z_\beta)$  from  $z$  he would incur a self-control cost  $S_\alpha := v(c + \alpha) - v(c)$ . Let  $D_\gamma := u(c + \gamma) - u(c)$  be the utility gain associated with the reward  $\gamma \in \{\alpha, \beta\}$ . Substituting  $u(c + \beta) + \delta W(z_0)$  for  $W(z_\beta)$  we can simplify the above inequality to establish that the DM will choose the smaller, earlier reward  $\alpha$  if  $S_\alpha > \delta D_\beta - D_\alpha$ .

Now consider the decision problem  $z'$ . From  $z'$  the agent chooses the reward  $\alpha$  if

$$u(c) + \delta u(c + \alpha) + \delta^2 W(z_0) > u(c) + \delta u(c) + \delta^2 W(z_\beta)$$

In this case, the agent experiences no self-control costs with either choice because the decision does not affect consumption in the current period. Substituting for  $W(z_\beta)$  the above inequality simplifies to  $0 > \delta D_\beta - D_\alpha$ . We observe a “preference reversal” if

$$S_\alpha > \delta D_\beta - D_\alpha > 0$$

When both of the inequalities above hold, the smaller, earlier reward is chosen if the reward can be consumed *immediately*, that is, in the decision period. But, if the decision is made earlier, the larger, later reward is chosen.

## 5. Measures of Preference for Commitment and Self-Control

In this section, we introduce measures that allow us to compare the preference for commitment and the self-control of decision-makers. These measures are based on similar concepts in Gul and Pesendorfer (2001). The versions here are weaker to facilitate the analysis of the applications considered in the next section.

To distinguish between differences in impatience and differences in preference for commitment (or self-control) we compare agents’ behavior in decision problems that involve no intertemporal trade-offs. The IHCP  $z$  is *intertemporally inconsequential* if every choice  $\mu \in z$  has the same marginal over continuation problems. Recall that for  $\mu \in \Delta(C \times Z)$ ,



$\mu^1$  denotes the marginal on  $C$  (current consumption) and  $\mu^2$  denotes the marginal on  $Z$  (continuation problem).

**Definition:**  $z$  is *intertemporally inconsequential (II)* if, for every  $\mu, \nu \in z$ ,  $\mu^2 = \nu^2$ . Let  $Z_{II}$  denote the set of all *intertemporally inconsequential IHCP's*.

The definition below presents a comparative measure of preference for commitment and of self-control.

**Definition:** The preference  $\succeq_1$  has *more instantaneous preference for commitment [self-control]* than  $\succeq_2$  if, for every  $z \in Z_{II}$ ,  $\succeq_2$  has preference for commitment [self-control] at  $z$  implies  $\succeq_1$  has preference for commitment [self-control] at  $z$ . The preferences  $\succeq_1$  and  $\succeq_2$  have the *same instantaneous preference for commitment [self-control]* if  $\succeq_1$  has more instantaneous preference commitment [self-control] than  $\succeq_2$  and  $\succeq_2$  has more instantaneous preference for commitment [self-control] than  $\succeq_1$ .

Our objective is to characterize these measures in terms of the representation  $(u, v, \delta)$ . This characterization assumes that the preferences are *regular*. A DSC preference  $\succeq$  with representation  $(u, v, \delta)$  is regular if  $v \neq \alpha u + \beta$  for any  $\alpha, \beta \in \mathbb{R}$ .

Consider two regular preferences  $\succeq_i$  with representations  $(u_i, v_i, \delta_i)$ ,  $i = 1, 2$ . Let  $z = \{\mu, \nu\} \in Z_{II}$  be such that  $E_\mu[u_1] > E_\nu[u_1]$ . Note that this implies  $E_\mu[u + \delta W] > E_\nu[u + \delta W]$  because  $\mu$  and  $\nu$  lead to the same marginal over continuation problems. Therefore, if  $\succeq_1$  has no preference for commitment at  $z$  then it must be that  $E_\mu[v_1] \geq E_\nu[v_1]$ . Now suppose that there are non-negative constants  $\alpha_u, \alpha_v, \beta_u, \beta_v$  such that

$$u_2 = \alpha_u u_1 + \alpha_v v_1$$

$$v_2 = \beta_u u_1 + \beta_v v_1$$

It follows that  $E_\mu[u_2] \geq E_\nu[u_2]$  and  $E_\mu[v_2] \geq E_\nu[v_2]$  and hence  $\succeq_2$  has no preference for commitment at  $x$ . Theorem 4 in Appendix C demonstrates that this condition is necessary and sufficient for  $\succeq_1$  to have more preference for commitment than  $\succeq_2$ .

An analogous result characterizes our measure of self-control. Consider  $z = \{\mu, \nu\} \in Z_{II}$  with  $E_\mu[u_1 + v_1] > E_\nu[u_1 + v_1]$ . Then,  $\mu$  is the optimal choice from  $z$ . If  $\succeq_1$  has no self-control at  $x$  then it must be that  $\mu$  is at least as tempting as  $\nu$  and hence  $E_\mu[v_1] \geq E_\nu[v_1]$ .

Now suppose that there are non-negative constants  $\alpha_u, \alpha_v, \beta_u, \beta_v$  such that

$$\begin{aligned} u_2 + v_2 &= \alpha_u(u_1 + v_1) + \alpha_v v_1 \\ v_2 &= \beta_u(u_1 + v_1) + \beta_v v_1 \end{aligned}$$

It follows that  $E_\mu[u_2 + v_2] \geq E_\nu[u_2 + v_2]$  and  $E_\mu[v_2] \geq E_\nu[v_2]$ . Therefore,  $\succeq_2$  has no self-control at  $z$ . Theorem 5 in Appendix C demonstrates that this condition is necessary and sufficient for  $\succeq_1$  to have more self-control than  $\succeq_2$ .

The following corollary analyzes situations where decision-makers either have the same preference for commitment and differ with respect to their self-control or have the same self-control and differ with respect to their preference for commitment. Since we only make instantaneous comparisons in this paper, henceforth, without risk of confusion we say “preference for commitment” and “self-control” rather than “instantaneous preference for commitment” and “instantaneous self-control”.

**Corollary 1:** *Let  $\succeq_i, i = 1, 2$  be regular DSC preferences with representation  $(u_i, v_i, \delta_i)$ . Then,*

(i)  *$\succeq_1$  and  $\succeq_2$  have the same preference for commitment and  $\succeq_1$  has more self-control than  $\succeq_2$  if and only if there exist  $\gamma \geq 1$  and  $\delta_2 \in (0, 1)$  such that  $(u_1, \gamma v_1, \delta_2)$  represents  $\succeq_2$ ;*

(ii)  *$\succeq_1$  and  $\succeq_2$  have the same self-control and  $\succeq_2$  has more preference for commitment than  $\succeq_1$  if and only if there exist  $\gamma \geq 1$  and  $\delta_2$  such that  $(u_1 + (1 - \gamma)v_1, \gamma v_1, \delta_2)$  is a representation of  $\succeq_2$ .*

**Proof:** See Appendix.

Part (i) of the Corollary says that keeping  $u$  constant and changing  $v$  to  $\gamma v$  for some  $\gamma > 1$  is equivalent to a decrease in self-control without changing preference for commitment. Part (ii) of the Corollary says that keeping  $u + v$  constant and changing  $v$  to  $\gamma v$  for some  $\gamma > 1$  is equivalent to increasing preference for commitment without changing self-control. We utilize this observation in our analysis of competitive equilibria.

## 6. Competitive Economies

In this section, we present examples of competitive economies with consumers who have DSC preferences. There are  $n$  consumers,  $i \in \{1, \dots, n\}$  each with a DSC preference  $(u_i, v_i, \delta_i)$ . We assume that each preference is regular, that is,  $v_i$  is not an affine transformation of  $u_i$ . This ensures that we can apply the results of Corollary 1 when using the comparative measures of preference for commitment and self-control.

We also assume the functions  $u_i$  and  $v_i$  are strictly increasing;  $v_i$  is convex and continuously differentiable;  $u_i + v_i$  is concave and continuously differentiable. The curvature assumptions imply that the temptation utility is risk-neutral or risk loving and the commitment utility is risk-neutral or risk averse. These assumptions ensure that the maximization problems below have concave objective functions. In terms of behavior, a risk loving temptation utility implies that the consumer is tempted by *instantaneous* gambles, that is, gambles that yield consumption in the current period. Since  $u_i + v_i$  is assumed to be concave, this temptation does not lead to risk-loving behavior.

There are  $L$  physical goods, indexed by  $l \in \{1, \dots, L\}$ ; the consumption in each period is contained in the compact set  $C$  where  $C := \{c \in \mathbb{R}_+^L \mid 0 \leq c_l \leq k\}$  for some  $k > 0$ .

### 6.1 Deterministic Exchange Economies

We first consider a deterministic exchange economy with complete markets. Consumers take prices as given and must choose a consumption vector each period subject to an inter-temporal budget constraint. Consumer  $i$  has endowment  $\omega_i = (\omega_{i1}, \dots, \omega_{it}, \dots)$ . Endowments are bounded away from zero, that is, there is an  $\varepsilon > 0$  such that  $\omega_{itl} > \varepsilon$  for all  $i, t, l$ . We denote with  $\omega = (\omega_1, \dots, \omega_n)$  the vector of endowments.

The sequence  $p = (p_1, \dots, p_t, \dots) \in (\mathbb{R}_+^L)^\infty$  denotes the period 1 prices of consumption. For a given price  $p$  we now define the decision problem of a consumer. Let  $b \geq 0$  denote the consumer's wealth at the start of period  $t$ . The consumer must choose  $(c, b')$  in the budget set

$$B_t(p, b) := \{(c, b') \mid cp_t + b' = b, c \in C, b' \geq 0\}$$

The corresponding IHCP is denoted by  $x_t(p, b)$  and defined recursively as

$$x_t(p, b) = \{(c, x_{t+1}(p, b')) \mid (c, b') \in B_t(p, b)\}$$

For a consumer with DSC preference  $(u_i, v_i, \delta_i)$ , the utility of the IHCP  $x_t(p, b)$  is:

$$\begin{aligned} W_i(x_t(p, b)) &= \max_{(c, b') \in B_t(p, b)} \{u_i(c) + v_i(c) + \delta_i W_i(x_{t+1}(p, b'))\} - \max_{(c, b') \in B_t(p, b)} v_i(c) \\ &= \max_{(c, b') \in B_t(p, b)} \{(u_i(c) + v_i(c) + \delta_i W_i(x_{t+1}(p, b')))\} - \max_{\{c \mid p_t c \leq b\}} v_i(c) \end{aligned}$$

Hence, given the prices  $p$  and wealth  $b_1$ , the consumer will choose a sequence  $(c_t, b_{t+1})$  that solves

$$\max_{(c_t, b_{t+1}) \in B_t(p, b_t)} u_i(c_t) + v_i(c_t) + \delta_i W_i(x_{t+1}(p, b_{t+1})) \quad (3)$$

Let  $\mathbf{c}_i = (c_{1i}, \dots, c_{ti}, \dots)$  denote consumer  $i$ 's consumption choices. Note that when a consumer chooses a feasible  $\mathbf{c}_i$ , the corresponding  $b_2, b_3, \dots$ , are determined uniquely. We say that  $\mathbf{c}_i$  is optimal for consumer  $i$  at prices  $p$  and wealth  $b_1$  if  $\mathbf{c}_i$  solves (3). The vector  $\mathbf{c} = (\mathbf{c}_1, \dots, \mathbf{c}_n)$  denotes an allocation for the economy.

For the special case of one physical good ( $L = 1$ ) the following first order condition must hold for all  $t \geq 1$  at an interior solution of (3):

$$u'_i(c_t) + v'_i(c_t) = \delta \frac{p_t}{p_{t+1}} (u'_i(c_{t+1}) + v'_i(c_{t+1}) - v'_i(b_{t+1}/p_{t+1})) \quad (4)$$

To get an intuition for equation (4) suppose the consumer reduces consumption in period  $t$  by a small amount to finance a corresponding increase in consumption in period  $t + 1$ . The left hand side represents the (temptation and commitment) utility loss due to the reduction in period  $t$  consumption. The term  $\delta \frac{p_t}{p_{t+1}} (u'_i(c_{t+1}) + v'_i(c_{t+1}))$  represents the (temptation and commitment) utility gain due to the increase in period  $t + 1$  consumption. The reduction in period  $t$  consumption increases the wealth in period  $t + 1$ . Therefore, the most tempting consumption choice in period  $t + 1$  changes as a result of the reduction in period  $t$  consumption. The term  $\delta \frac{p_t}{p_{t+1}} v'_i(b_{t+1}/p_{t+1})$  represents the increase in the temptation utility of the most tempting choice in period  $t + 1$ . This term has a negative sign since it increases the cost of self-control. It is this last term that distinguishes DSC preferences from standard time separable utility functions.

**Definition:** The pair  $(p, \mathbf{c})$  is an equilibrium for the economy  $((u_i, v_i, \delta_i), \omega_i)_{i=1}^n$  if (1) for all  $i$ ,  $\mathbf{c}_i$  is optimal at prices  $p$  and wealth  $p \cdot \omega_i$ ; (2)  $\sum_{i=1}^n c_i = \sum_{i=1}^n \omega_i$ .

**Proposition 1:** *An equilibrium for the deterministic exchange economy exists.*

**Proof:** see Appendix.

To examine the welfare properties of equilibria we must define the appropriate notion of efficiency. Recall that a consumer's utility depends not only on his consumption but also on the possible consumption vectors each period. Therefore, the definition of admissible interventions for a social planner and hence the definition of Pareto efficiency must specify not only the feasible allocations of consumption but also the ways in which the social planner can restrict the set of feasible choices for consumers. If, for example, the social planner can impose arbitrary restrictions on choice sets then he can improve welfare simply by restricting the consumers' choice sets to the singleton set containing only the equilibrium allocation.

We assume that the planner may re-allocate resources or change prices but cannot put additional restrictions on the feasible choices of consumers. Hence, the consumers' decision problem in period  $t$  must be of the form  $x_t(p, b)$  for some price vector  $p$  and some wealth  $b$ . Since we permit a limited set of interventions for the planner we obtain a weak notion of Pareto efficiency. We will show that competitive equilibria with a representative DSC consumer may fail even this weak notion of efficiency.

**Definition:** *The pair  $(p, \mathbf{c})$  is admissible if for all  $i$   $\mathbf{c}_i$  is optimal at prices  $p$ , wealth  $p \cdot \mathbf{c}_i$  and if  $\sum_{i=1}^n \mathbf{c}_i \leq \sum_{i=1}^n \omega_i$ . An admissible pair  $(p, \mathbf{c})$  is Pareto optimal if there is no admissible  $(p', \mathbf{c}')$  with  $W_i(x_1(p', p' \cdot \mathbf{c}'_i)) > W_i(x_1(p, p \cdot \mathbf{c}_i))$  for all  $i$ .*

The following example demonstrates that a competitive equilibrium may not be Pareto optimal.

**Example:** Consider an economy with two physical goods and a representative consumer. The utility function is given by  $u(c_1, c_2) = \log c_1 - \lambda c_1 + c_2$  and  $v(c_1, c_2) = \lambda c_1$ . The consumption set is  $C = [0, 2]^2$ , the endowment is  $(1, 1)$  in every period and  $\lambda \in (0, \frac{1}{2})$ . Consider a policy where the planner confiscates (and destroys)  $\epsilon > 0$  units of good 1 in every period. For  $\epsilon$  small, this policy increases the consumer's welfare. A more detailed calculations can be found Appendix D. Here we provide the intuition. Recall that the cost of self-control in period 1 is the equilibrium value of the temptation utility  $\lambda(1 - \epsilon)$

minus the maximal temptation utility. The maximal temptation utility is achieved by converting all wealth into current consumption of good 1. In equilibrium, the reduction in the endowment of good 1 has two effects: first it decreases the equilibrium consumption of good 1, second it increases the price of good 1. The first effect reduces welfare while the second effect increases welfare by reducing the number of units of good 1 the consumer can afford in period 1 and hence reducing the cost of self-control. For the particular functional forms we have chosen, the second effect dominates the first. Note that the government could also levy a tax on good 1 and use the proceeds to buy (and destroy) good 1. This policy is equivalent to the confiscation policy and hence the example shows that such a tax policy may be welfare improving.

## 6.2 Steady State Equilibria

Next, we examine steady state equilibria of a particular class of deterministic exchange economies. We assume that there is one physical good, i.e.,  $L = 1$  and that consumers share a common discount factor  $\delta$ . In addition, all consumers have the same preference for commitment but may differ in their self-control. By Corollary 1 this implies that there is  $(u, v, \delta)$  and  $(\lambda_1, \dots, \lambda_n)$  with  $\lambda_i > 0$  such that  $i$ 's preferences are given by  $(u, \lambda_i v, \delta)$ . If  $\lambda_i > \lambda_j$  then  $j$  has more self-control than  $i$ .

In a steady-state equilibrium each consumer's consumption is constant over time and prices are characterized by a fixed interest rate  $r$ . Clearly, for an economy to have a steady state equilibrium, the aggregate endowment must be constant over time. Let  $\bar{\omega} \in C$  be the aggregate endowment in any period and let  $p^r := (1, (1+r)^{-1}, \dots, (1+r)^{-t}, \dots)$  be the price sequence corresponding to the interest rate  $r$ . The interest rate  $r$  and the consumption vector  $(c_1, \dots, c_I) \in C^I$  are a steady state equilibrium for the aggregate endowment  $\bar{\omega} \in C$  if (1)  $\sum_{i=1}^I c_i = \bar{\omega}$  and (2)  $(p^r, \mathbf{c})$  with  $\mathbf{c}_i = (c_i, \dots, c_i, \dots)$  is an equilibrium for the economy with endowments  $\mathbf{c}$ .

Proposition 2 shows that to each aggregate endowment there is a unique steady-state equilibrium. Hence, steady-state allocations are uniquely determined by the aggregate resources and the DSC preferences.

**Proposition 2:** Suppose that  $u + \lambda_i v$  is strictly concave for all  $i$  and  $\lim_{c \rightarrow 0} (u'(c) + v'(c)) \rightarrow \infty$ . Then, there is a unique steady-state equilibrium  $(r, c_1, \dots, c_I)$  for every aggregate endowment  $\bar{\omega}$ . Moreover,  $c_i > c_j$  iff  $\lambda_i < \lambda_j$ , that is, consumer  $i$ 's steady state consumption is higher than consumer  $j$ 's if and only if  $i$  has more self-control than  $j$ .

**Proof:** See Appendix.

Note that the conditions of Proposition 2 guarantee an interior solution to the consumer's maximization problem. Therefore, equation (4) implies that if  $(r, c_1, \dots, c_I)$  is a steady state equilibrium then:

$$(\delta(1+r) - 1)(u'_i(c_i) + v'_i(c_i)) = \delta(1+r)v'_i(c_i(1+r)/r) \quad (4')$$

where  $c_i(1+r)/r$  is the (constant) wealth of consumer  $i$  in terms of consumption. Since  $v' > 0$  we must have  $\delta(1+r) > 1$ . The concavity of  $u_i + v_i$  and the convexity of  $v_i$  imply that for each interest rate  $r$  with  $\delta(1+r) > 1$  there is a unique  $c_i$  consistent with equation (4'). The uniqueness of the steady state then follows from the fact that a higher interest rate implies a higher steady state consumption for all consumers.

For standard consumers with  $v_i \equiv 0$  the steady state interest rate satisfies  $\delta(1+r) = 1$ . In that case, any consumption  $c_i$  satisfies equation (4') and hence there are many possible steady state equilibria for any aggregate endowment. By contrast, aggregate endowment and the distribution of preferences uniquely determine the steady state in an exchange economy with DSC preferences.

As an example, let  $u(c) = \log c$  and  $v(c) = c$ . In that case, (4') simplifies to

$$1 + \lambda_i c = \delta(1+r)$$

Setting  $\gamma_i = 1/\lambda_i$ , the unique solution to these equations and the aggregate resource constraint is

$$c = \frac{\bar{\omega}\gamma_i}{\sum_j \gamma_j}$$

and

$$1 + r = \frac{\bar{\omega} + \sum_j \gamma_j}{\delta \sum_j \gamma_j}$$

Hence, a mean preserving spread in the  $\gamma_i$ 's will leave the interest rate unchanged but result in greater inequality in the steady state. On the other hand, an increase in  $\bar{\omega}$ , the aggregate endowment, will leave dispersion unchanged but lead to an increase in the steady state interest rate.

DSC preferences have the feature that the marginal rate of substitution between consumption in period  $t$  and  $t + 1$  depends on the maximal consumption in period  $t + 1$  and hence on the consumer's wealth in that period. Therefore, DSC preferences have similar implications for time preference as the utility functions introduced by Uzawa (1968) and Epstein and Hynes (1983). In these models, the marginal rate of substitution between periods  $t$  and  $t + 1$  may depend on the consumption in all periods  $t' \geq t$  and hence implicitly on the wealth in period  $t + 1$ .

### 6.3 Borrowing constraints

In this section we illustrate the effect of borrowing constraints on equilibrium consumption and prices. Consider the following example of a deterministic exchange economy with one physical good. There are two consumers with identical preferences  $(u, \lambda v, \delta)$  where

$$u(c) = c - \frac{\lambda c^2}{2}, \quad v(c) = \frac{\lambda c^2}{2}$$

with  $\delta > \alpha$ ,  $\lambda \in (0, \frac{\delta - \alpha}{3\delta})$  and  $\alpha = \frac{\sqrt{7}-1}{3}$ . The aggregate endowment is 3 every period and

$$\omega_{1t} = 3 - \omega_{2t} = \begin{cases} 2 & \text{if } t \text{ odd} \\ 1 & \text{if } t \text{ even} \end{cases}$$

When there are no borrowing constraints this economy has a unique equilibrium in which both agents consume  $\frac{3}{2}$  in all but the first period. To see this, let  $\bar{c}_{it}$  denote the maximally feasible consumption of consumer  $i$  in period  $t$ . The first order condition for consumer  $i$ 's maximization problem (equation (4), substituting  $\bar{c}_{it+1}$  for  $b_{it+1}/p_{t+1}$ ) implies

$$1 = \frac{p_t}{p_{t+1}} \delta (1 - \lambda \bar{c}_{it+1}) \tag{4''}$$

In equilibrium, the first order condition (4'') can only be satisfied for both consumers if  $\bar{c}_{1t+1} = \bar{c}_{2t+1}$  for all  $t \geq 1$ . Since  $\bar{c}_{it}$  is equal to the consumer's wealth in terms of period  $t$  consumption, it follows that both consumers must have the same wealth and



the same consumption in all but the first period. Period 1 consumption is such that the wealth of both individuals is the same starting in period 2. Equilibrium therefore predicts consumption smoothing by both consumers. The equilibrium prices satisfy

$$\frac{p_t}{p_{t+1}} = \frac{2}{1 + \delta - \sqrt{(1 - \delta)^2 + 6\delta\lambda}}$$

Now assume that consumers face a borrowing constraint. In particular, the maximum amount each consumer can borrow is next period's endowment. Let  $s$  denote a consumer's savings at the end of the previous period. Consumer  $i$ 's budget set is denoted  $B_{it}(p, s)$  where

$$B_{it}(p, s) = \{(c, s') \mid sp_{t-1} + p_t\omega_{it} = p_t s' + p_t c, -s' \leq \omega_{it+1}, c \in C\}$$

The first order condition of the consumer's optimization problem is again given by (4''). The borrowing constraint affects the maximally feasible consumption  $\bar{c}_{it}$ . Suppose consumers choose  $s = 0$  in all periods and consume their endowment. Then  $\bar{c}_{it} = 3$  (the sum of current and next period's endowment). If

$$\frac{p_t}{p_{t+1}} = \frac{1}{\delta(1 - 3\lambda)}$$

then this choice satisfies the first order condition (4''). Hence,  $\mathbf{c}_i = \omega_i$  and the price sequence  $p$  with  $p_t = (\delta(1 - 3\lambda))^{t-1}$  is an equilibrium of the economy with borrowing constraints. Note that in equilibrium the borrowing constraint is not binding.

This example illustrates the consumers' desire to smooth the maximally feasible consumption. Smoothing of the maximally feasible consumption is achieved by refraining from consumption smoothing. Since  $v$  is convex, smoothing temptation economizes on the cost of self-control. Note that we have chosen a linear  $u + v$  and hence have eliminated the consumer's incentive to smooth actual consumption. In general, both motives for consumption smoothing are present when consumers have DSC preferences.

The interest rate  $r = \frac{p_t}{p_{t+1}} - 1$  in the equilibrium with the borrowing constraint is lower than the interest rate when there is no borrowing constraint. To see why this is the case, note that in the equilibrium with no borrowing constraints the maximally feasible consumption  $b_{t+1}$  is greater than 3 for both consumers. The introduction of a borrowing

constraint therefore reduces  $\bar{c}_{it+1}$  and hence the cost of self-control. Equation (4'') then implies that  $p_t/p_{t+1}$  must adjust downwards to offset this effect.

In the equilibrium above, consumption tracks income because borrowing opportunities are much greater in low income periods than in high income periods. The individual refrains from shifting funds to low income periods to avoid increasing the cost of self-control. With a different borrowing constraint one could obtain the opposite result: that is, consumption could be larger in periods where the individual has low endowment. For example, if the individual may borrow 1 unit independent of his future endowment, then equilibrium consumption in periods of high endowment could be lower than in periods of low endowment.

#### 6.4 Stochastic, Representative Agent Economy

In this subsection, we analyze a simple example of a stochastic representative consumer economy (Lucas (1978)). There is one consumer who owns a productive asset that yields a dividend in each period. Dividends are identically and independently distributed across time and denoted by the random variable  $d$ . Let  $D$  denote the realized value of  $d$ . We assume that  $d$  has mean 1, variance  $\sigma^2$  and support  $[0, 2]$ .

Let  $s$  denote the asset holdings of the representative consumer at beginning of the current. Then, the consumer observes  $D$ , the realization of dividends in the current period, and chooses consumption  $c \in C := [0, 4]$ . Given the price of the asset  $p$ , the choice of  $c$  determines  $s'$ , the consumer's holding of the asset for next period. The price of consumption is normalized to 1. The price of the productive asset depends on the dividend realization and is described by the function  $p : [0, 2] \rightarrow \mathbb{R}_+$ . Let

$$B(s, D) := \{(c, s') \mid p(D)s + Ds = p(D)s' + c\}$$

denote the consumer's budget set given the asset holding  $s$  and the realization of the dividend  $D$ .

It is easy to see that the decision problem of the representative consumer defines an IHCP for each initial asset holding  $s$  and each initial value of the dividend  $D$ . Let  $y(s, D)$  denote this IHCP. Similarly, let  $y(s, d)$  denote the IHCP confronting the representative consumer before the current period dividend is realized.

The utility of  $y(s, D)$  for a consumer with the DSC preference  $(u, v, \delta)$  satisfies

$$W(y(s, D)) = \max_{(c, s') \in B(s, D)} \{u(c) + v(c) + \delta W(y(s', d))\} - v([D + p(D)]s)$$

Hence, a consumer with the DSC preference  $(u, v, \delta)$  chooses consumption to solve

$$\max_{(c, s') \in B(s, D)} u(c) + v(c) + \delta W(y(s', d)) \quad (5)$$

The first order conditions for a solution to (5) is

$$\begin{aligned} p(D)(u'(c_t) + v'(c_t)) &= \delta E [(u'(\tilde{c}_{t+1}) + v'(\tilde{c}_{t+1}))(p(d) + d)] \\ &\quad - \delta E [v'(s_{t+1}(p(d) + d))(p(d) + d)] \end{aligned} \quad (6)$$

In a competitive equilibrium  $c_t = D$ ,  $s_{t+1} = 1$ , and  $\tilde{c}_{t+1} = d$ .

To illustrate the effect of temptation on asset prices we consider the following example:

$$u(c) = c - \frac{\lambda c^2}{2}, \quad v(c) = \frac{\lambda c^2}{2}, \quad 0 < \lambda < \frac{1}{2}$$

By Corollary 1, an increase in  $\lambda$  corresponds to an increase in the consumer's preference for commitment while keeping self-control constant. Note that for  $\lambda = 0$  this consumer is a standard risk-neutral decision-maker. For  $\lambda > 0$  the consumer remains risk-neutral with respect to *instantaneous consumption gambles*, that is, gambles that replace current consumption  $c$  with a lottery of consumptions with mean  $c$ . However, the consumer is *risk averse* with respect to investments that affect future wealth. To illustrate this, we compute the risk premium, that is, the difference in the expected return between a risky and a risk-free asset, for this economy.

Equation (6) applied to this example together with the equilibrium conditions yield a constant price  $p$  that satisfies

$$p = \delta(1 + p - 2\lambda p - \lambda p^2 - \lambda(\sigma^2 + 1))$$

Let  $\frac{1}{1+r}$  denote the price of a risk-free asset that pays off one unit of consumption in the next period, irrespective of the dividend realization. In equilibrium, (see Appendix D for details) the risk-free rate must satisfy

$$\frac{1}{1+r} = \delta(1 - \lambda(p + 1))$$

Let  $R$  be the expected return on the productive asset, that is,  $1 + R = E \left[ \frac{d+p}{p} \right] = \frac{1+p}{p}$ . Appropriate substitution (see Appendix D) yields the following expression for the equity premium:

$$R - r = \lambda \sigma^2 \frac{\delta(1+r)}{p}$$

The equity premium is positive for  $\lambda > 0$  and increasing in  $\lambda$  (see Appendix D for detailed calculations).

The utility of an individual with DSC preferences depends both on actual and maximally feasible consumption (wealth). When  $u + v$  is concave and  $v$  is convex, there are two sources of risk aversion. First, as with standard concave utility functions, the consumer is averse to consumption risk. Second, the consumer is averse to risk in wealth because the cost of self-control is convex. Our example illustrates the second source of risk aversion. In the example, the consumer is risk-neutral with respect to consumption ( $u + v$  is linear) but risk averse with respect to realizations of future wealth. Since investment in a risky asset implies that future wealth is uncertain, risky assets must offer a risk premium.

When choosing among lotteries that promise immediate (consumption) rewards the DSC consumer exhibits less risk aversion than when choosing among assets that promise risky future returns. In the former case, self-control costs are sunk and hence do not affect the consumer's choice whereas in the latter case self-control costs add to the consumer's risk aversion.

## 7. Sustainable Debt

In this section we show that for an agent with DSC preferences, incentive compatible debt contracts are feasible even in an environment where the only punishment for default is exclusion from future borrowing. Hence, the agent may save funds at the market interest rate even after default. This restriction on feasible punishments after default is particularly relevant in the case of sovereign debt. With standard preferences (i.e. no preference for commitment), Bulow and Rogoff (1989) show that in this environment there is no incentive compatible contract that allows the individual to borrow.

We assume that there are no investment opportunities that offer commitment. Thus, the consumer always has the option of exchanging his savings for current consumption.

This assumption may be justified by allowing *collateralized* loans after default. To see how this works, suppose that in period 1 the consumer invests in a contract that offers a return of 1 unit of consumption in period 3. In period 2 the consumer is able to use that contract as collateral for a loan on current consumption. Under this hypothesis there are incentive compatible contracts that allow individuals with DSC preferences to borrow. In the following we provide an example of such a contract.

Let  $\omega = (0, 0, 4, 0, 0, 4, 0, 0, 4, \dots)$  be the agent's endowment. The agent borrows and lends at a fixed interest rate  $r$ . A generic debt contract is denoted  $(\beta_1, \beta_2, \beta_3)$  with the understanding that  $\beta_j$  is the required outstanding balance at the end of any period  $3t + j$ . If, at the end of any period, the agent's balance is not at the required level, then he is excluded from borrowing in all future periods. The individual may always invest funds at a rate  $r > 0$ , even after default.

The contract  $(\beta_1, \beta_2, \beta_3)$  is incentive compatible if for any feasible debt level in period  $t$ , there is an optimal plan in which the consumer does not default. Simple calculations establish

$$\begin{aligned}\beta_1 &= (1+r)\beta_3 + c_{3n+1} \\ \beta_2 &= (1+r)\beta_1 + c_{3n+2} \\ \beta_3 &= (1+r)\beta_2 + c_{3n} - 4\end{aligned}$$

for  $n = 1, 2, \dots$ . Since consumption is nonnegative, the above equations imply that if there is any borrowing (i.e.,  $\beta_j > 0$  for some  $j$ ) then  $\beta_2 \geq \beta_1 \geq \beta_3$ . Hence, borrowing occurs if and only if  $\beta_2 > 0$  and the maximal level of debt is  $\beta_2$ . Finally, the above three equations imply

$$\frac{c_{3n+2}}{(1+r)^2} + \frac{c_{3n+1}}{(1+r)} + c_{3n} = 4 - \left[ (1+r) - \frac{1}{(1+r)^2} \right] \beta_2 \quad (7)$$

for all  $n \geq 1$ . Under any contract with  $\beta_2 > 0$ , a standard agent with no preference for commitment has an incentive to default. To see this, suppose, instead of repaying the debt the individual "deposits"  $\frac{c_{3n+1}}{1+r} + \frac{c_{3n+2}}{(1+r)^2}$  into a savings account. By (7) this is feasible and yields strictly more than  $c_3$  units of consumption for period 3 whenever  $\beta_2 > 0$ . This argument is a special case of the argument given by Bulow and Rogoff (1989) to demonstrate that without direct penalties there can be no borrowing.

However, if the agent has a preference for commitment, then incentive compatible borrowing is possible. Let  $(u, v, \delta)$  be the agent's utility function, where

$$u(c) = \begin{cases} 2c & \text{if } c \leq 1 \\ 1 + c & \text{if } c > 1 \end{cases}$$

$$v(c) = \lambda c$$

Assume that  $\delta$  is close enough to 1, so that  $\alpha := \frac{1-\delta^2}{\delta^3} + \frac{1-\delta^4}{\delta^4} < \frac{2\delta^2-1}{2\delta^2+\delta}$ . Let  $\underline{\lambda} := \frac{\alpha}{1-\alpha}$  and  $\bar{\lambda} := \frac{2\delta^2-1}{1+\delta}$ .

**Proposition 3:** *If  $\lambda \in (\underline{\lambda}, \bar{\lambda})$  then for all  $r \leq \frac{1-\delta}{\delta}$ , the debt contract  $\beta_1 = 1, \beta_2 = 2 + r, \beta_3 = 0$  is incentive compatible.*

**Proof:** First, consider the optimal program for an agent who cannot borrow starting in period 3 (the period when he will be tempted to default). Note that it cannot be optimal to consume more than 1 in periods other than  $3t$ . To see this recall that  $(1+r)\delta \leq 1$  and consumption beyond 1 has a marginal utility of 1 in each period. Hence the incentive to avoid the self-control penalty at period  $3t$  makes it optimal to consume no more than 1 in all periods but  $3t$ . On the other hand, consuming less than 1 in any period cannot be optimal: suppose that the individual consumes less than 1 in period 5. Consider an increase in period 5 consumption financed by a decrease in period 3 consumption. The marginal utility of consumption in period 5 is 2, discounted by two periods yields  $2\delta^2$ . The marginal utility of consumption in period 3 is 1. In addition, the transfer of resources from period 3 to period 5 increases self-control costs in periods 3 and 4. The marginal increase in self-control costs is  $\lambda(1+\delta)$ . Since  $\lambda < \bar{\lambda}$ , the increase of period 5 consumption increases payoff. A similar argument holds for period 4 and hence the individual will consume exactly one unit in periods  $3t+1, 3t+2, t \geq 1$ . Therefore, the optimal utility starting from any period  $3t$  is:

$$W_d^3 = \frac{1}{1-\delta^3} \left[ 5 - \frac{1+\lambda}{1+r} - \frac{1+\lambda}{(1+r)^2} + \left( 2 - \frac{\lambda}{1+r} \right) \delta + 2\delta^2 \right]$$

Now, consider the consumer who does not default. In period  $3t+2$ , the continuation utility of the plan is

$$W_p^3 = \frac{1}{1-\delta^3} \left[ 5 - (1+\lambda)(1+r) - (1+\lambda)(1+r)^2 + 2\delta + 2\delta^2 \right]$$

Straightforward calculations, using the facts  $r \leq \frac{1-\delta}{\delta}$  and  $\lambda > \underline{\lambda}$  establish that  $W_p^3 > W_d^3$ . Therefore, the individual has no incentive to default in any period  $3t + 2$ . But this is the period with the highest incentive to default. Hence, the debt contract is incentive compatible. *Q.E.D.*

The idea behind Proposition 3 extends to other utility functions: compared to the savings program, the borrowing program leads to a lower self-control costs in periods  $3t + 1$ . The reason is that an agent who enters period 5 with a debt  $(1+r)\beta_1$  is extended additional credit equal to  $\beta_2 - (1+r)\beta_1 > 0$ . These funds are not available to the agent in period 4 and therefore he does not suffer the self-control costs associated with “transferring” them to period 5. Unlike savings, “credit-worthiness” is not an asset that can be used as collateral. Therefore, the commitment offered by the debt contract cannot be undone in the open market.

## 8. Conclusion

The starting point of the time-inconsistency literature is a non-recursive preference  $\succeq_1$  over consumption streams  $\{c_t\}_{t=1}^\infty$ . Then, it is assumed that the preference  $\succeq_\tau$  over consumption streams starting at time  $\tau$  is the same as  $\succeq_1$ . Since  $\succeq_1$  is not recursive this implies that the conditional preference relation induced by  $\succeq_1$  on consumption streams starting at date  $\tau$  is not the same as  $\succeq_\tau$ . That is, there exists some  $\{c_1, \dots, c_{\tau-1}, c_\tau, \dots\}$  and  $\{c_1, \dots, c_{\tau-1}, \hat{c}_\tau, \dots\}$  such that

$$\begin{aligned} \{c_1, \dots, c_{\tau-1}, c_\tau, \dots\} &\succeq_1 \{c_1, \dots, c_{\tau-1}, \hat{c}_\tau, \dots\} \\ \text{and } \{\hat{c}_\tau, \dots\} &\succ_\tau \{c_\tau, \dots\} \end{aligned}$$

This “reversal” of preference is called time-inconsistency.

The time-inconsistency approach postulates that at each  $\tau$  the agent is a different “player” and behaves in a manner that maximizes  $\succeq_\tau$  given the predicted behavior of his subsequent selves. The predicted consumption paths are typically the subgame perfect Nash outcomes of the resulting game.<sup>12</sup>

The multiplicity of subgame perfect Nash equilibria implies that a given decision problem may not have a unique payoff associated with it. From a game-theoretic perspective,

this is not surprising; we almost never expect all equilibria of a given game to yield the same payoff for a particular agent. In a multi-person context, subgame perfect Nash equilibrium is meant to capture a rest point of the player’s expectations and strategizing. Then, the multiplicity of such equilibria is a reflection of the fact that no single person controls the underlying forces that might lead to the particular rest point.

But this multiplicity is more difficult to understand within the context of a single person decision problem, even if the person in question is dynamically inconsistent. In that case, the game-theoretic argument for multiplicity loses much of its force since it should be straightforward to re-negotiate one’s self out of an unattractive continuation equilibrium. And, foresight of this renegotiation would lead to the unraveling of the original plan.<sup>13</sup> More generally, the notion of (subgame perfect) Nash equilibrium is a tool for the analysis of *non-cooperative* behavior, and its appropriateness often rests on the implicit assumption of “independent” behavior and absence of communication. Therefore, analyzing the interaction between the agent at time  $\tau$  and his slightly modified self at time  $\tau + 1$  as the Nash equilibrium of a game may not be appropriate.<sup>14</sup>

In contrast, the approach of Gul and Pesendorfer (2001) and the current paper is to take a single preference, not over consumption, but over a class of decision problems and to permit a strict preference for a smaller set of options. In our approach, the preference for commitment arises from a desire to avoid temptation rather than from a change in preference. As demonstrated by Theorem 2, multiplicity does not arise in the preference for commitment approach. Every  $(u, v, \delta)$  corresponds to a unique preference over decision problems and all optimal plans yield the same payoff.

Within the time-inconsistency approach, the only formulation of self-control entails using multiplicity of equilibria to construct equilibria in which the decision-maker sustains a desirable plan by threatening *himself* with less desirable behavior after a deviation. Self-control is therefore not a property of the agent’s preference but of the selected equilibrium and a theory of self-control requires a theory of equilibrium selection. By contrast, the preference for commitment approach identifies self-control as a property of the agent’s preference. This allows us to identify parameters of the agent’s utility function that measure the amount of self-control this agent has.



*Dept. of Economics, Princeton University, Princeton, NJ 08540, U.S.A.;*  
*fgul@princeton.edu; www.princeton.edu/fgul,*

*and*

*Dept. of Economics, Princeton University, Princeton, NJ 08540, U.S.A.;*  
*pesendor@princeton.edu; www.princeton.edu/pesendor.*

## 9. Appendix A: Infinite Horizon Consumption Problems

Our treatment of dynamic decision problems is similar to the “descriptive approach” in Kreps and Porteus (1978) extended to an infinite horizon.

Let  $X$  be any metric space. The set  $\mathcal{K}(X)$ , of all non-empty compact subsets of  $X$  (endowed with the *Hausdorff* metric) is itself a metric space. If  $X$  is compact then so is  $\mathcal{K}(X)$  (see Brown and Percy (1995) p. 222). Let  $\Delta(X)$  denote the set of all measures on the Borel  $\sigma$ -algebra of  $X$ . We endow  $\Delta(X)$  with the weak topology. If  $X$  is compact then  $\Delta(X)$  is also compact and metrizable (with the Prohorov metric) (see Parthasarathy (1970)). For any metric space  $X$ , we use  $\mathcal{B}(X)$  to denote the Borel  $\sigma$ -algebra of  $X$ .

Given any sequence of metric spaces  $X_t$ , we endow  $\times_{t=1}^{\infty} X_t$  with the product topology. This topology is also metrizable and  $\times_{t=1}^{\infty} X_t$  is compact if each  $X_t$  is compact (Royden (1968) pp. 152, 166).

Let  $C$  denote the compact metric space of possible consumptions in each period. Let  $Z_1 := \mathcal{K}(\Delta(C))$ . An element of  $Z_1$  is a one period consumption problem. Each choice  $\mu_1 \in z_1 \in Z_1$  is a probability measure on  $C$ . For  $t > 1$ , define  $Z_t$  inductively as  $Z_t := \mathcal{K}(\Delta(C \times Z_{t-1}))$ . Thus, each  $z_t$  in  $Z_t$  is a  $t$  period consumption problem. An element  $\mu_t \in z_t$  is a probability measure on  $(C, Z_{t-1})$ , that is, a probability measure on consumption in the current period and  $t - 1$  period consumption problems.

Let  $Z^* := \times_{t=1}^{\infty} Z_t$ . The set of infinite horizon consumption problems (IHCP) are those elements of  $Z^*$  that are consistent, that is, for  $z = \{z_t\}_{t=1}^{\infty} \in Z^*$  the  $t - 1$  period consumption problem induced by  $z_t$  is equal to  $z_{t-1}$ . To be more precise, let  $G_1 : C \times Z_1 \rightarrow C$  be given by

$$G_1(c, z_1) := c$$

and let  $F_1 : \Delta(C \times Z_1) \rightarrow \Delta(C)$ ,  $\bar{F}_1 : \mathcal{K}(\Delta(C \times Z_1)) \rightarrow \mathcal{K}(\Delta(C))$  be defined as follows:

$$F_1(\mu_2)(E) := \mu_2(G_1^{-1}(E))$$

$$\bar{F}_1(z_2) := \{F_1(\mu_2) \mid \mu_2 \in z_2\}$$

for  $E$  in the Borel  $\sigma$ -algebra of  $C$ . Thus,  $F_1(\mu_2)$  is the probability measure over current consumption induced by  $\mu_2 \in z_2$  and  $\bar{F}_1(z_2)$  is the one period decision problem induced by  $z_2$ . Proceeding inductively, we define  $G_t : C \times Z_t \rightarrow C \times Z_{t-1}$  by

$$G_t(c, z_t) := (c, \bar{F}_{t-1}(z_t))$$

and  $F_t : \Delta(C \times Z_t) \rightarrow \Delta(C \times Z_{t-1})$ ,  $\bar{F}_t : \mathcal{K}(\Delta(C \times Z_t)) \rightarrow \mathcal{K}(\Delta(C \times Z_{t-1}))$  by

$$F_t(\mu_{t+1})(E) := \mu_{t+1}(G_t^{-1}(E))$$

$$\bar{F}_t(z_{t+1}) := \{F_t(\mu_{t+1}) \mid \mu_{t+1} \in z_{t+1}\}$$

for  $E \in \mathcal{B}(C \times Z_{t-1})$ . Then,  $\bar{F}_t(z_t)$  is the  $t - 1$  period decision problem induced by  $z_t$ . Finally, we define

$$Z := \{\{z_t\}_{t=1}^\infty \in Z^* \mid z_{t-1} = \bar{F}_{t-1}(z_t) \forall t > 1\}$$

to be the set of all IHCP's.

**Lemma 1:** *Let  $X, Y$  be compact metric spaces and  $g : X \rightarrow Y$  be a continuous function. Then, (i)  $\bar{g} : \mathcal{K}(X) \rightarrow \mathcal{K}(Y)$  defined by  $\bar{g}(A) = \{g(a) \mid a \in A\}$  is also continuous. (ii) If  $g$  is a bijection, then it is a homeomorphism.*

**Proof:** Part (i) follows from exercise X, p. 222 in Brown and Percy (1995). Part (ii) follows from Theorem 8.37, p. 207 in Brown and Percy (1995). *Q.E.D.*

Note that  $F_1$  is continuous and hence by Lemma 1(i), so is  $\bar{F}_1$ . Then, an inductive argument establishes that  $F_t, \bar{F}_t$  are continuous for all  $t \geq 1$ . It follows that  $Z$ , the set of all IHCP's is compact.

As described, every IHCP is an infinite sequence of finite decision problems. Alternatively, an IHCP can also be viewed as a set of options, each of which results in a probability measures over current consumption and IHCP's that describe the consumer's

situation next period. Hence, the recursive view identifies an IHCP as an element of  $\mathcal{K}(\Delta(C \times Z))$ . Indeed, there is a natural mapping from  $Z$  to  $\mathcal{K}(\Delta(C \times Z))$ .

**Theorem A1:** *There exists a homeomorphism  $f : Z \rightarrow \mathcal{K}(\Delta(C \times Z))$ .*

To illustrate how  $f$  identifies recursive IHCPs with IHCPs, consider the problem  $\{z_t\}_{t=1}^\infty \in Z$  and assume that each choice is deterministic, that is; each  $\mu_t \in z_t$  yields a (certain) period 1 consumption,  $c_t$ , and a (certain) continuation problem  $z_{t-1}$ . Consider a sequence  $\{\mu_t\}_{t=1}^\infty$  such that  $\mu_t \in z_t$  and  $\mu_{t-1} = F_{t-1}(\mu_t)$  for all  $t$ . Let  $(c_t, z_{t-1})$  denote the element of  $C \times Z_{t-1}$  that occurs with probability 1 according to  $\mu_t$ . Since  $\{\mu_t\}_{t=1}^\infty$  is consistent ( $\mu_{t-1} = F_{t-1}(\mu_t)$ ) it follows that  $c_t = c_1$  for all  $t$ ; that is, the period 1 consumption induced by  $\mu_t$  is the same as the period 1 consumption induced by  $\mu_1$ . Moreover, the sequence of continuation problems  $z' := \{z_{t-1}\}_{t=2}^\infty$  is itself consistent. Hence, we can identify  $\{\mu_t\}_{t=1}^\infty$  with the unique element in  $\Delta(C \times Z)$  that puts probability 1 on  $(c, z')$ . Repeating this procedure for every consistent sequence  $\{\mu_t\}$  such that  $\mu_t \in z_t$  for all  $t$  yields  $f(z)$ .

The construction of the set of IHCP's shares the following common structure with several recent contributions. Let  $Y_0$  be a set with some property  $\alpha$  and let  $\mathcal{C}$  be an operator that associates with each set that has property  $\alpha$  a new set that also has property  $\alpha$ . Let  $Y_1 := \mathcal{C}(Y_0)$  and  $Y_{t+1} := \mathcal{C}(Y_0 \times Y_t)$  for  $t \geq 1$  and let  $Y^* = \times_{t=0}^\infty Y_t$ . The goal is to identify a subset  $Y$  of  $Y^*$  that is homeomorphic to the set  $\mathcal{C}(Y_0 \times Y^*)$ .

In Mertens and Zamir (1985), Epstein and Zin (1989), Brandenburger and Dekel (1993),  $Y_0$  is a topological space with some property (for example, a Polish space in Brandenburger and Dekel) and  $\mathcal{C}(Y_0)$  is a set of measures on  $Y_0$ . In Epstein and Wang (1996),  $Y_0$  is compact Hausdorff and  $\mathcal{C}(Y_0)$  are preferences over a set of functions from  $Y_0$  to  $[0, 1]$ . In Epstein and Peters (1999)  $\mathcal{C}(Y_0)$  is a set of upper hemi-continuous functions on  $Y_0$  and in Mariotti and Piccione (1999),  $\mathcal{C}(Y_0)$  is the set of all non-empty compact subsets of the compact set  $Y_0$ . In our case,  $Y_0$  is a compact metric space and  $\mathcal{C}(Y_0)$  is the set of all non-empty, compact subsets of probability measures on the Borel  $\sigma$ - algebra of  $Y_0$ .

**Definition:** *Let  $X$  be a compact metric space. Let  $A^n \in \mathcal{K}(X)$  for all  $n$ . The closed limit inferior of the sequence  $A^n$  (denoted  $\underline{L}A^n$ ) is the set of all  $a \in X$  such that  $a = \lim a^n$*

for some sequence  $\{a^n\}$  such that  $a^n \in A^n$  for every  $n$ . The closed limit superior of  $A^n$  (denoted  $\bar{L}A^n$ ) is the set of all  $a \in X$  such that  $a = \lim a^{n_j}$  for some subsequence  $\{a^{n_j}\}$  such that  $a^{n_j} \in A^{n_j}$  for every  $n_j$ . The set  $LA^n$  is the pointwise limit of  $A^n$  if  $LA^n = \underline{L}A^n = \bar{L}A^n$ .

**Lemma 2:** Let  $X$  be a compact metric space. The sequence  $A^n \in \mathcal{K}(X)$  converges to  $A$  iff  $A = LA^n$ .

**Proof:** The lemma follows from exercise X, p.132 in Brown and Pearcy (1995). *Q.E.D.*

**Definition:** Let  $\Upsilon_1 := \Delta(C)$  and for  $t > 1$  let  $\Upsilon_t := \Delta(C, Z_1, \dots, Z_{t-1})$ . The sequence of probability measures  $\{\hat{\mu}_t\} \in \times_{k=1}^{\infty} \Upsilon_t$  is Kolomogorov consistent if  $\text{marg}_{C, Z_1, \dots, Z_{t-1}} \hat{\mu}_{t+1} = \hat{\mu}_t$  for all  $t \geq 1$ . Let  $\Upsilon^{kc}$  denote all the set of all Kolomogorov consistent sequences in  $\times_{t=1}^{\infty} \Upsilon_t$ .

**Lemma 3:** For every  $\{\hat{\mu}_t\} \in \Upsilon^{kc}$  there exists a unique  $\mu \in \Delta(C \times Z^*)$  such that  $\text{marg}_C \mu = \hat{\mu}_1$  and  $\text{marg}_{C, \dots, Z_t} \mu = \hat{\mu}_t$  for all  $t \geq 1$ . The mapping  $\psi : \Upsilon^{kc} \rightarrow \Delta(C \times Z^*)$  that associates this  $\mu$  with the corresponding  $\{\hat{\mu}_t\}$  is a homeomorphism.

**Proof:** The first assertion is Kolmogorov's Existence Theorem [Dellacherie and Meyer (1978), p. 73]. Since every compact space is complete and separable, the second assertion follows from Lemma 1 in Brandenburger and Dekel (1993).

**Definition:** Let  $D_t := \{(z_1, \dots, z_t) \in \times_{n=1}^t Z_t \mid z_k = \bar{F}_k(z_{k+1}) \forall k = 1, \dots, t-1\}$ ,  $M^c := \{\{\mu_t\} \in \times_{t=1}^{\infty} \Delta(C) \times \Delta(C \times Z_t) \mid F_t(\mu_{t+1}) = \mu_t \forall t \geq 1\}$  and  $\Upsilon^c := \{\{\hat{\mu}_t\} \in \Upsilon^{kc} \mid \hat{\mu}_{t+1}(C \times D_t) = 1 \forall t \geq 1\}$ .

**Lemma 4:** For every  $\{\mu_t\} \in M^c$  there exists a unique  $\{\hat{\mu}_t\} \in \Upsilon^c$  such that  $\hat{\mu}_1 = \mu_1$  and  $\text{marg}_{C, Z_{t-1}} \hat{\mu}_t = \mu_t$  for all  $t \geq 2$ . The mapping  $\phi : M^c \rightarrow \Upsilon^c$  that associates this  $\{\mu_t\}$  with the corresponding  $\{\hat{\mu}_t\}$  is a homeomorphism.

**Proof:** Let  $m_0$  be the identity function on  $C$  and let  $m_1$  be the identity function on  $C \times Z_1$ . For  $t \geq 2$ , define  $m_t : C \times Z_t \rightarrow C \times (\times_{k=1}^t Z_k)$  as follows:  $m_t(c, z_t) = (\hat{c}, \hat{z}_1 \dots \hat{z}_t)$  iff  $\hat{c} = c, \hat{z}_t = z_t$  and  $\hat{z}_{k-1} = \bar{F}_{k-1}(\hat{z}_k)$  for all  $k = 2, \dots, t$ . Note that  $m_t$  is one-to-one for all  $t$ . Also,  $m_t(C \times Z_t) = C \times D_t$ . Let  $\pi_0$  and  $\pi_1$  be the identity mappings on  $C$  and  $C \times Z_1$  respectively. For  $t \geq 2$ , let  $\pi_t(c, \dots, z_t) = (c, z_t)$  for all  $(c, \dots, z_t) \in C \times D_t$ .

Clearly,  $\pi_t$  is continuous for all  $t$ . Since  $C \times D_t$  is a compact set  $\pi_t$  can be extended to all of  $C \times (\times_{k=1}^t Z_k)$  in a continuous manner. Hence,  $\pi_t$  is continuous function on  $C \times (\times_{k=1}^t Z_k)$  and its restriction to  $C \times D_t$  is the inverse of  $m_t : C \times Z_t \rightarrow m_t(C \times Z_t)$ . Since  $\bar{F}_t$  is continuous for all  $t$ , so is  $m_t$ .

For  $\{\mu_t\} \in M^c$ , define  $\{\hat{\mu}_t\}$  by  $\hat{\mu}_t(E) := \mu_t(m^{-1}(E))$  for all  $E \in \mathcal{B}(C \times (\times_{k=1}^t Z_k))$ . Clearly, the  $\{\hat{\mu}_t\}$  defined in this manner is the unique element in  $\Upsilon^c$  such that  $\hat{\mu}_1 = \mu_1$  and  $\text{marg}_{C, Z_{t-1}} \hat{\mu}_t = \mu_t$  for all  $t \geq 1$ . Define  $\phi(\{\mu_t\})$  to be this unique  $\{\hat{\mu}_t\}$  and note that  $\phi$  is one-to-one. Pick any  $\{\hat{\mu}_t\}$  in  $\Upsilon^c$ . Define  $\{\mu_t\}$  as follows  $\mu_1 = \hat{\mu}_1$  and  $\mu_t(E) := \hat{\mu}_t(\pi_{t-1}^{-1}(E))$  for all  $E \in \mathcal{B}(C \times Z_{t-1})$ . Note that  $\phi(\{\mu_t\}) = \hat{\mu}_t$  hence,  $\phi$  is a bijection. Observe that the  $t$ 'th element of  $\phi(\{\mu_t\})$  depends only on  $\mu_t$ . Hence, without risk of confusion we write  $\phi_t(\mu_t)$  to denote this element. Note that for any continuous real-valued function  $\hat{h}$  on  $C \times (\times_{k=1}^t Z_k)$  and  $h$  on  $C \times Z_t$ ,  $\int \hat{h} d\phi_t(\mu_t) = \int \hat{h} \circ m_t d\mu_t$  and  $\int h d\phi_t^{-1}(\hat{\mu}_t) = \int h \circ \pi_t d\hat{\mu}_t$ . Hence, the continuity of  $\phi$  and  $\phi^{-1}$  follows from the continuity of  $m_t$  and  $\pi_t$  for all  $t$ . *Q.E.D.*

**Lemma 5:**  $\psi(\Upsilon^c) = \{\mu \in \Delta(C \times Z^*) \mid \mu(C \times Z) = 1\}$ .

**Proof:** Let  $\Gamma_t = C \times D_t \times \times_{k=t+1}^{\infty} Z_k$  for all  $t \geq 1$  and  $\mu = \psi(\{\hat{\mu}_t\})$ . Observe that  $\mu(\Gamma_t) = \hat{\mu}_t(C \times D_t) = 1 \forall t$  if  $\{\hat{\mu}_t\} \in \Upsilon^c$ . Hence  $\mu(C \times Z) = \mu(\bigcap_{t \geq 1} \Gamma_t) = \lim \mu(\Gamma_t) = 1$ . Conversely, if  $\mu(C \times Z) = 1$  then  $\mu(\Gamma_t) = 1 \forall t$  and hence there is a corresponding  $\{\hat{\mu}_t\} \in \Upsilon^c$ . *Q.E.D.*

**Lemma 6:** Let  $\xi(z) := \{\{\mu_t\} \in M^c \mid \mu_t \in z_t \forall t \geq 1\}$ . Then  $\xi : Z \rightarrow \mathcal{K}(M^c)$  is a homeomorphism and  $\{\mu_t\} \in \xi(z)$  iff  $\mu_t \in z_t$  for all  $t$ .

**Proof:** Step 1: Let  $z = \{z_t\} \in Z$ ,  $\mu_\tau \in z_\tau$ . Then, there exists  $\{\nu_t\} \in \xi(z)$  such that  $\nu_\tau = \mu_\tau$ .

Proof of Step 1: Let  $\nu_\tau = \mu_\tau$ . For  $k = 1, \dots, \tau - 1$ , define  $\nu_{\tau-k}$  inductively as  $\nu_{\tau-k} := F_{\tau-k+1}(\mu_{\tau-k+1})$ . Similarly, define  $\nu_{\tau+k}$ , for  $k \geq 1$  inductively by picking any  $\nu_{\tau+k} \in \bar{F}_{\tau+k}^{-1}(z_{\tau+k-1}) \cap z_{\tau+k}$ . The  $\{\nu_t\}$  constructed in this fashion has the desired properties concluding the proof of step 1.

By Step 1,  $\xi(z) \neq \emptyset$ . To see that  $\xi(z)$  is compact, take any sequence  $\{\mu_t^{n_j}\} \in \xi(z)$  for  $z = (z_1, z_2, \dots)$ . We can use the diagonal method to find a subsequence  $\{\mu_t^{n_j}\}$  such that  $\mu_t^{n_j}$  converges to some  $\mu_t$  for every  $t$ . Since each  $z_t$  is compact,  $\mu_t \in z_t$  for all  $t$ . Then, the continuity of  $F_t$  for all  $t$ , implies  $\{\mu_t\}$  is in  $M^c$  and therefore in  $\xi(z)$ . So,  $\xi(z)$

is compact. Suppose  $z \neq z'$  for some  $z, z' \in Z$ . Without loss of generality, assume there exists some  $\tau$  and  $\mu_\tau$  such that  $\mu_\tau \in z_\tau \setminus z'_\tau$ . By Step 1, we obtain  $\{\nu_t\} \in \xi(z)$  such that  $\nu_\tau = \mu_\tau$ . Clearly,  $\{\nu_t\} \in \xi(z) \setminus \xi(z')$ . Therefore,  $\xi$  is one-to-one. Take any  $\bar{z} \in \mathcal{K}(M^c)$ . Define  $z_\tau := \{\mu_\tau \mid \bar{\mu}_\tau = \mu_\tau \text{ for some } \{\bar{\mu}_t\} \in \bar{z}\}$ . Let  $z = (z_1, z_2, \dots)$ . Since  $\bar{z}$  is compact we have  $z \in Z$ . Clearly,  $\bar{z} \subset \xi(z)$ . Using the compactness of  $\bar{z}$  again and the continuity of each  $F_t$  implies  $\xi(z) \subset \bar{z}$ . So,  $\xi$  is onto.

To prove that  $\xi$  is continuous, let  $z^n$  converge to  $z$ . By Lemma 2, this is equivalent to  $Lz_t^n = z_t$  for all  $t$ . We need to show that  $\xi(z^n)$  converges to  $\xi(z)$ . Take any convergent sequence  $\{\mu_t^n\}$  such that  $\{\mu_t^n\} \in \xi(z^n)$  for all  $n$ . Then,  $\lim \mu_t^n \in z_t$  for all  $t$  and therefore  $\lim\{\mu_t^n\} \in \xi(z)$ . Let  $\{\mu_t\} \in \xi(z)$ . Since  $Lz_t^n = z_t$ , there exists  $\mu_t^n$  converging to  $\mu_t$  such that  $\mu_t^n \in z_t^n$  for all  $n$ . By step 1, for each  $\mu_\tau^n$  we can construct  $\{\nu_t\}(\tau) \in M^c$  such that  $\nu_\tau(\tau) = \mu_\tau^n$ . Since  $F_t$  is continuous for all  $t$ ,  $\nu_k(\tau)$  converges to  $\mu_k$  for all  $k \leq \tau$ . Consequently,  $\{\nu_t\}(n)$  converges to  $\{\mu_t\}$  as  $n$  goes to infinity. Hence,  $L\xi(z^n) = \xi(z)$ . Again by Lemma 2, this implies  $\xi(z^n)$  converges to  $\xi(z)$  and hence  $\xi$  is continuous.

To conclude the proof recall that  $Z$  is compact and  $\xi$  is a continuous bijection. It follows from Lemma 1(ii) that  $\xi$  is a homeomorphism. *Q.E.D.*

**Proof of Theorem A1:** Note that by Lemmas 3–5,  $\psi \circ \phi$  is a homeomorphism from  $M^c$  to  $\Delta(C \times Z)$ . Hence, by Lemma 1(i), the function  $\zeta : \mathcal{K}(M^c) \rightarrow \mathcal{K}(\Delta(C \times Z))$  defined by  $\zeta(A) := \psi \circ \phi(A)$  for all  $A \in \mathcal{K}(M^c)$  is also a homeomorphism. Then, by Lemma 6,  $\zeta \circ \xi$  is the desired homeomorphism from  $Z$  to  $\mathcal{K}(\Delta(C \times Z))$ . *Q.E.D.*

## 10. Appendix B: Proof of Theorem 1

**Lemma 7 (A Fixed-Point Theorem):** *If  $B$  is a closed subset of a Banach space with norm  $\|\cdot\|$  and  $T : B \rightarrow B$  is a contraction mapping (i.e., for some integer  $m$  and scalar  $\alpha \in (0, 1)$ ,  $\|T^m(W) - T^m(W')\| \leq \alpha \|W - W'\|$  for all  $W, W' \in B$ ), then there is a unique  $W^* \in B$  such that  $T(W^*) = W^*$ .*

**Proof:** See [Bertsekas and Shreve (1978), p. 55] who note that the theorem in Ortega and Rheinolt (1970) can be generalized to Banach spaces.

**Lemma 8:** Let  $u : C \rightarrow \mathbb{R}$ ,  $v : C \rightarrow \mathbb{R}$  be continuous and  $\delta \in (0, 1)$ . There is a unique continuous function  $W : Z \rightarrow \mathbb{R}$  such that

$$W(z) = \max_{\mu \in z} \left\{ \int (u(c) + v(c) + \delta W(z')) d\mu(c, z') \right\} - \max_{\nu \in z} \int v(c) d\nu(c, z') \quad (8)$$

for all  $z \in Z$ .

**Proof:** Let  $\mathcal{W}$  be the Banach space of all continuous, real-valued functions on  $Z$  (endowed with the sup norm). The operator  $T : \mathcal{W} \rightarrow \mathcal{W}$ , where

$$TW(z) = \max_{\mu \in z} \left\{ \int (u(c) + v(c) + \delta W(z')) d\mu(c, z') \right\} - \max_{\nu \in z} \int v(c) d\nu(c, z')$$

is well-defined and is a contraction mapping. Hence, by Lemma 7, there exists a unique  $W$  such that  $T(W) = W$ . Hence,  $W$  satisfies (8). *Q.E.D.*

By Lemma 8, for any continuous  $u, v, \delta$ , there exists a unique continuous  $W$  that satisfies (8). It is straightforward to verify that Axioms 1 – 8 hold for any binary relation represented by a continuous function  $W$  that satisfies (8).

In the remainder of the proof we show that if  $\succeq$  is non-degenerate and satisfies Axioms 1 – 8 then the desired representation exists. It is easy to show that if  $\succeq$  satisfies Axioms 4, 6 and 7 then it also satisfies the following stronger version of the independence axiom (see footnote 5):

**Axiom 4\*:**  $x \succ y$ ,  $\alpha \in (0, 1)$  implies  $\alpha x + (1 - \alpha)z \succ \alpha y + (1 - \alpha)z$ .

Theorem 1 of Gul and Pesendorfer (2001) establishes that  $\succeq$  satisfies Axioms 1 – 3 and 4\* if and only if there exist linear and continuous functions  $U, V$  such that the function  $W$  defined by

$$W(z) := \max_{\mu \in z} \{U(\mu) + V(\mu)\} - \max_{\nu \in z} V(\nu) \quad (9)$$

represents  $\succeq$ . To complete the proof we will show that there exist continuous functions  $u, v$  and  $\delta \in (0, 1)$  such that for all  $z \in Z$ ,

$$\begin{aligned} U(\mu) &= \int u(c) + \delta W(z) d\mu(c, z) \\ V(\mu) &= \int v(c) d\mu(c, z) \end{aligned}$$

**Lemma 9:** *There exists a continuous  $u : C \rightarrow \mathbb{R}$ ,  $\delta \in (0, 1)$ ,  $\gamma \in \mathbb{R}$  such that  $U(\mu) = \int (u(c) + \delta W(z)) d\mu(c, z) + \gamma$  for all  $\mu \in \Delta(C \times Z)$ .*

**Proof:** *Step 1:* There are continuous  $u : C \rightarrow \mathbb{R}$  and  $\bar{W} : Z \rightarrow \mathbb{R}$  such that  $U(\mu) = \int (u(c) + \bar{W}(z)) d\mu(c, z)$  for all  $\mu \in \Delta$ .

*Proof:* Since  $U$  is linear and continuous there exists a continuous  $\bar{u} : C \times Z \rightarrow \mathbb{R}$  such that  $U(\mu) = \int \bar{u}(c, z) d\mu(c, z)$ . By Axiom 5,  $U(.5(\bar{c}, \bar{z}) + .5(c, z)) = U(.5(\bar{c}, z) + .5(c, \bar{z}))$  and therefore,  $\bar{u}(c, z) = \bar{u}(c, \bar{z}) + \bar{u}(\bar{c}, z) - \bar{u}(\bar{c}, \bar{z})$ . Hence,

$$\begin{aligned} U(\mu) &= \int \bar{u}(c, z) d\mu(c, z) \\ &= \int \bar{u}(c, \bar{z}) d\mu(c, z) + \int \bar{u}(\bar{c}, z) d\mu(c, z) - \int \bar{u}(\bar{c}, \bar{z}) d\mu(c, z) \end{aligned}$$

Setting  $u := \bar{u}(\cdot, \bar{z}) - \bar{u}(\bar{c}, \bar{z})$  and  $\bar{W} := \bar{u}(\bar{c}, \cdot)$  gives the desired result.

*Step 2:* There exist  $\delta > 0$ ,  $\gamma \in \mathbb{R}$  such that  $\bar{W}(z) = \delta W(z) + \gamma$  for all  $z \in Z$ .

*Proof:* Define  $K := \max_Z W(z)$ ,  $k := \min_Z W(z)$ ,  $\bar{K} := \max_Z \bar{W}(z)$ ,  $\bar{k} := \min_Z \bar{W}(z)$ . By non-degeneracy,  $U$  is not constant. It follows that  $W$  is not constant. Axioms 6 implies that  $\bar{W}(x) \geq \bar{W}(y)$  if and only if  $W(x) \geq W(y)$ . Hence,  $\bar{W}$  is not constant. Therefore,  $\bar{K} > \bar{k}$ ,  $K > k$ . To establish the desired conclusion we will show that

$$\bar{W}(z) := \frac{\bar{K}k - K\bar{k}}{(K - k)(\bar{K} - \bar{k})} + \frac{\bar{K} - \bar{k}}{K - k} W(z) \quad (10)$$

for all  $z \in Z$ . For any  $z \in Z$  there exists a unique  $\alpha$  such that

$$\bar{W}(z) = \alpha \bar{K} + (1 - \alpha) \bar{k} \quad (11)$$

Let  $z^*$  maximize  $\bar{W}$  and  $z_*$  minimize it. It follows that

$$\begin{aligned} W(\{\alpha(\bar{c}, z^*) + (1 - \alpha)(\bar{c}, z_*)\}) &= \alpha W(\{(\bar{c}, z^*)\}) + (1 - \alpha)W(\{(\bar{c}, z_*)\}) \\ &= u(\bar{c}) + \alpha \bar{K} + (1 - \alpha) \bar{k} \\ &= u(\bar{c}) + \bar{W}(z) \\ &= W(\{(c, z)\}) \end{aligned}$$



(The first equality follows from linearity of  $W$ , the second equality follows from (9) and Step 1, the third equality follows from (11), and the last equality follows from (9).) By Axiom 7 we conclude that

$$W(\{(\bar{c}, z)\}) = W(\{(\bar{c}, \alpha z^* + (1 - \alpha)z_*)\})$$

and by Axiom 6

$$W(z) = W(\alpha z^* + (1 - \alpha)z_*)$$

Linearity of  $W$  together with the fact that  $\bar{W}(x) \geq \bar{W}(y)$  iff  $W(x) \geq W(y)$  implies

$$W(z) = \alpha K + (1 - \alpha)k \tag{12}$$

Solving (12) for  $\alpha$ , substituting the result into (11) and re-arranging terms then yields (10) and proves step 2.

*Step 3:  $\delta < 1$ .*

*Proof:* Let  $z^c$  be the unique  $z \in Z$  with the property that  $z^c = \{(c, z^c)\}$ . Let  $z$  be such that  $W(z) \neq W(z^c)$ . Let  $y^1 = \{(c, z)\}$  and define  $y^n$  inductively as  $y^n = \{(c, y^{n-1})\}$ . Then  $y^n$  converges to  $z^c$ . Hence, by continuity  $W(y^n) - W(z^c)$  must converge to zero. But

$$W(y^n) - W(z^c) = \delta^n(W(z) - W(z^c))$$

Since  $W(z) - W(z^c) \neq 0$  it follows that  $\delta < 1$ .

*Q.E.D.*

Let  $U' = U - \frac{\gamma}{1-\delta}$  and  $W' = W - \frac{\gamma}{1-\delta}$ . Then,  $W', U'$  are continuous and linear with

$$W'(z) := \max_{\mu \in z} \{U'(\mu) + V(\mu)\} - \max_{\nu \in z} V(\nu)$$

Moreover,  $W'$  represents  $\succeq$  and

$$U'(\mu) = \int (u(c) + \delta W'(z)) d\mu(c, z)$$

Therefore, without loss of generality, we can set  $\gamma = 0$  in Lemma 9.

**Lemma 10:** Assume there exists  $\bar{\mu}$  and  $\underline{\mu}$  such that  $U(\bar{\mu}) + V(\bar{\mu}) - U(\underline{\mu}) - V(\underline{\mu}) > 0 > V(\bar{\mu}) - V(\underline{\mu})$ . Then, there is a continuous  $v : C \rightarrow \mathbb{R}$  such that  $V(\mu) = \int v(c)d\mu(c, z)$  for all  $\mu \in \Delta$ .

**Proof:** Fix  $\mu^2 \in \Delta(Z)$  and define  $V^* : \Delta(C) \rightarrow \mathbb{R}$  by

$$V^*(\mu^1) := V(\mu^1 \times \mu^2)$$

Take any  $\nu \in \Delta$  and let  $\mu$  be the product measure  $\nu^1 \times \mu^2$ . By continuity, there exists  $\alpha > 0$  small enough so that

$$U(\alpha\nu + (1 - \alpha)\underline{\mu}) + V(\alpha\nu + (1 - \alpha)\underline{\mu}) < U(\bar{\mu}) + V(\bar{\mu})$$

$$U(\alpha\mu + (1 - \alpha)\underline{\mu}) + V(\alpha\mu + (1 - \alpha)\underline{\mu}) < U(\bar{\mu}) + V(\bar{\mu})$$

$$V(\alpha\nu + (1 - \alpha)\underline{\mu}) > V(\bar{\mu})$$

$$V(\alpha\mu + (1 - \alpha)\underline{\mu}) > V(\bar{\mu})$$

Axiom 8 implies that  $W(\{\alpha\nu + (1 - \alpha)\underline{\mu}, \bar{\mu}\}) = W(\{\alpha\mu + (1 - \alpha)\underline{\mu}, \bar{\mu}\})$ . It now follows from (9) that  $V(\alpha\nu + (1 - \alpha)\underline{\mu}) = V(\alpha\mu + (1 - \alpha)\underline{\mu})$ . Since  $V$  is linear, we have  $V(\nu) = V(\mu) = V^*(\nu^1)$ . Since  $V$  is continuous and linear, so is  $V^*$ . Hence there is a  $v : C \rightarrow \mathbb{R}$  such that  $V^*(\nu^1) = \int v(c)d\nu^1(c)$  as desired. *Q.E.D.*

To complete to proof, we show that the conclusion of Lemma 10 holds in all cases. By non-degeneracy  $U$  is not constant. If  $V$  is constant the conclusion of Lemma 10 holds trivially. So, we assume that neither  $U$  nor  $V$  are constant.

Suppose  $V = \alpha U + \beta$  for some  $\alpha, \beta \in \mathbb{R}$ . Since  $V$  is not constant we conclude  $\alpha \neq 0$ . If  $\alpha > 0$ , replace  $V$  with  $V' = 0$ . Then,  $W(z) := \max_{\mu \in z} \{U(\mu) + V'(\mu)\} - \max_{\nu \in z} V'(\nu)$  and the conclusion of Lemma 10 holds. Non-degeneracy rules out  $\alpha \leq -1$ . If  $\alpha \in (-1, 0)$  or if  $V$  is not an affine transformation of  $U$  then  $V$  is not a positive affine transformation of  $U + V$ . Hence, the preferences represented by  $V$  and  $U + V$  are different and non-trivial (i.e. neither  $V$  nor  $U + V$  is constant). Therefore, there exists  $\bar{\nu}, \underline{\nu}$  such that either  $U(\bar{\nu}) + V(\bar{\nu}) \geq U(\underline{\nu}) + V(\underline{\nu})$  and  $V(\bar{\nu}) < U(\underline{\nu})$  or  $U(\bar{\nu}) + V(\bar{\nu}) > U(\underline{\nu}) + V(\underline{\nu})$  and  $U(\bar{\nu}) \leq U(\underline{\nu})$ . In either case, since neither  $U$  nor  $V$  is constant, we can use the linearity of  $U$  and  $V$  to find  $\bar{\mu}$  close to  $\bar{\nu}$  and  $\underline{\mu}$  close to  $\underline{\nu}$  for which all of the above inequalities are strict and apply Lemma 10. *Q.E.D.*

### 10.1 Proof of Theorem 3

A DSC preference  $\succeq$  represented by  $(u, v, \delta)$  has the property that  $U := u + \delta W$  is not constant. Furthermore, since  $\succeq$  has a preference for commitment at some  $z \in Z$  it follows that  $v$  is not constant. Consequently, there exists no  $\alpha, \beta \in \mathbb{R}$  such that  $v = \alpha U + \beta$ . Then, we may apply the proof of Theorem 4 in Gul and Pesendorfer (2001) to conclude that if  $(u, v, \delta)$  and  $(u', v', \delta')$  both represent  $\succeq$  then  $u' + \delta' W' = \alpha(u + \delta W) + \beta_u$ ,  $v' = \alpha v + \beta_v$  and therefore  $W' = \alpha W + \beta_u$  for some  $\alpha > 0, \beta_u, \beta_v \in \mathbb{R}$ . Since  $u$  is not constant this implies that  $\delta' = \delta$  and  $u' = \alpha u + (1 - \delta)\beta_u$ . The proof of the converse is straightforward and therefore omitted. Q.E.D.

## 11. Appendix C: Measures

**Theorem 4:** *Let  $\succeq_i, i = 1, 2$  be a regular DSC preferences with representation  $(u_i, v_i, \delta_i)$ . Then,  $\succeq_2$  has more instantaneous preference for commitment than  $\succeq_1$  if and only if there exist a non-singular, non-negative matrix  $\Theta'$  and  $\beta' \in \mathbb{R}^2$  such that*

$$\begin{pmatrix} u_1 \\ v_1 \end{pmatrix} = \Theta' \cdot \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} + \beta'$$

**Proof:** Let  $\succeq^*$  be a binary relation on  $\mathcal{A} := \mathcal{K}(\Delta(C))$ . Gul and Pesendorfer (2001) show that  $\succeq^*$  satisfies Axioms 1 – 3 and 4\* if and only if there exists  $(U^*, V^*)$  such that

$$W^*(A) := \max_{\mu^1 \in A} \{U^*(\mu^1) + V^*(\mu^1)\} - \max_{\mu^1 \in A} V^*(\mu^1) \quad (13)$$

represents  $\succeq^*$ . Then, they define comparative measures of preference for commitment, self-control as in section 4, but without the restriction to intertemporally inconsequential decision problems. If there is  $W^*, (U^*, V^*)$  satisfying (13) such that  $W^*$  represents  $\succeq^*$  we say  $(U^*, V^*)$  represents  $\succeq^*$ . The preference  $\succeq^*$  with representation  $(U^*, V^*)$  is regular\* if  $V^*$  is not an affine transformation of  $U^*$ . Theorem 8 of Gul and Pesendorfer (2001)

establishes that if  $\succeq_i^*$  with representations  $(U_i^*, V_i^*), i = 1, 2$  are regular\* preferences then  $\succeq_2^*$  has more preference for commitment than  $\succeq_1^*$  if and only if

$$\begin{aligned} U_1^* &= \alpha_u U_2^* + (1 - \alpha_u) V_2^* \\ V_1^* &= \alpha_v U_2^* + (1 - \alpha_v) V_2^* \end{aligned}$$

for  $\alpha_u, \alpha_v \in [0, 1]$ . We note that any binary relation  $\succeq$  on  $Z$  that satisfies Axioms 1 – 8 induces the following binary relation  $\succeq^*$  on  $K(\Delta(C))$ :  $A \succeq^* B$  iff  $x \succeq y$  for some  $x, y \in Z_{II}$  such that  $A := \{\mu^1 \mid \mu \in x\}, B := \{\mu^1 \mid \mu \in y\}$  and  $\mu^2 = \hat{\mu}^2$  for all  $\mu \in x, \hat{\mu} \in y$ . The preference  $\succeq^*$  is well-defined since it can be represented as in Theorem 2 for some  $(u, v, \delta)$ . Define  $U^*, V^*$  by  $U^*(\mu^1) := E_{\mu^1}[u], V^*(\mu) := E_{\mu^1}[v]$  and define  $W^*$  as in (13). It is easy to verify that  $W^*$  represents  $\succeq^*$  and since  $\succeq$  is regular, it follows that  $V^*$  is not an affine transformation of  $U^*$  and hence  $\succeq^*$  is regular\*. Also, if  $\succeq_1^*, \succeq_2^*$  are two preference relations induced respectively by  $\succeq_1, \succeq_2$ , then  $\succeq_2$  has more instantaneous preference for commitment than  $\succeq_1$  iff  $\succeq_2^*$  has more preference for commitment than  $\succeq_1^*$ . Therefore we may apply Theorem 8 in Gul and Pesendorfer (2001) and the uniqueness result of section 2 (Theorem 3) to yield the desired result (that  $\Theta$  is nonsingular follows from the regularity of  $\succeq_1$ ). Q.E.D.

**Theorem 5:** *Let  $\succeq_i, i = 1, 2$  be a regular DSC preference with representation  $(u_i, v_i, \delta_i)$ . Then,  $\succeq_1$  has more instantaneous self-control than  $\succeq_2$  if and only if there exist a non-singular, non-negative matrix  $\Theta'$  and  $\beta' \in \mathbb{R}^2$  such that*

$$\begin{pmatrix} u_2 + v_2 \\ v_2 \end{pmatrix} = \Theta' \cdot \begin{pmatrix} u_1 + v_2 \\ v_1 \end{pmatrix} + \beta'$$

**Proof:** We note that Gul and Pesendorfer (2001) provide a definition of comparative self-control analogous to one in section 4 but without the restriction to intertemporally inconsequential decision problems. In Theorem 9 they show that if  $\succeq_i^*$  with representations  $(U_i^*, V_i^*), i = 1, 2$  are regular\* preferences then  $\succeq_1^*$  has more self-control than  $\succeq_2^*$  if and only if

$$\begin{aligned} U_2^* + V_2^* &= \alpha_u (U_1^* + V_1^*) + (1 - \alpha_u) V_1^* \\ V_2^* &= \alpha_v (U_1^* + V_1^*) + (1 - \alpha_v) V_1^* \end{aligned}$$

for some  $\alpha_u, \alpha_v \in [0, 1]$ . Define  $\succeq_i^*$ ,  $U_i^*$  and  $V_i^*$  for  $i = 1, 2$  as in the proof of Theorem 4. Again, it is easy to verify that  $\succeq_1$  has more instantaneous self-control than  $\succeq_2$  iff  $\succeq_1^*$  has more self control than  $\succeq_2^*$ . Therefore we may apply Theorem 9 in Gul and Pesendorfer (2001) and Theorem 3 of section 2 to yield the desired result. *Q.E.D.*

**Proof of Corollary 1:** The “if” parts of both statements are straightforward and omitted.

By Theorem 4, if  $\succeq_1$  has more preference for commitment than  $\succeq_2$  and  $(u_i, v_i, \delta)$  is a representation of  $\succeq_i$  for  $i = 1, 2$  there is a non-singular, non-negative matrix  $\Theta$  and  $\beta \in \mathbb{R}^2$  such that

$$\begin{pmatrix} u_1 \\ v_1 \end{pmatrix} = \Theta \cdot \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} + \beta \quad (14)$$

Similarly, by Theorem 5, if  $\succeq_1$  has more self-control than  $\succeq_2$  and  $(u_i, v_i, \delta)$  is a representation of  $\succeq_i$  for  $i = 1, 2$  there is a non-singular, non-negative matrix  $\Theta'$  and  $\beta' \in \mathbb{R}^2$  such that

$$\begin{pmatrix} u_2 + v_2 \\ v_2 \end{pmatrix} = \Theta' \cdot \begin{pmatrix} u_1 + v_2 \\ v_1 \end{pmatrix} + \beta'$$

Suppose  $\succeq_1$  and  $\succeq_2$  have the same preference for commitment and  $\succeq_1$  has more self-control than  $\succeq_2$ . Without loss of generality we can choose  $(u_1, v_1)$  such that  $\beta = (0, 0)$  in (14). That is, for some non-singular, non-negative matrix  $\Theta$

$$\begin{pmatrix} u_1 \\ v_1 \end{pmatrix} = \Theta \cdot \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} \quad (15)$$

Similarly, reversing the roles of  $(u_1, v_1)$  and  $(u_2, v_2)$  in (14) yields a non-singular, non-negative  $\hat{\Theta}$  and  $\hat{\beta} \in \mathbb{R}^2$  such that

$$\begin{pmatrix} u_2 \\ v_2 \end{pmatrix} = \hat{\Theta} \cdot \begin{pmatrix} u_1 \\ v_1 \end{pmatrix} + \hat{\beta} \quad (16)$$

Equation (15) implies that  $\hat{\beta} = 0$  and  $\hat{\Theta} = \Theta^{-1}$ . But since both  $\hat{\Theta}$  and  $\Theta$  are non-negative, this implies

$$\hat{\Theta} = \begin{pmatrix} \alpha & 0 \\ 0 & \gamma \end{pmatrix}, \Theta = \begin{pmatrix} \frac{1}{\alpha} & 0 \\ 0 & \frac{1}{\gamma} \end{pmatrix} \quad (17)$$

for some  $\alpha > 0, \gamma > 0$ . Again, without loss of generality, we assume  $\alpha = 1$ . To conclude the proof, we show that  $\gamma \geq 1$ . Since  $\succeq_1$  has more self-control than  $\succeq_2$ , the regularity

of  $\succeq_2$ , equations (14) – (17) and the fact that  $\hat{\beta} = 0$  imply that for some non-negative, non-singular  $\tilde{\Theta}$  and  $\tilde{\beta}$ ,

$$\begin{pmatrix} u_2 + v_2 \\ v_2 \end{pmatrix} = \tilde{\Theta} \cdot \begin{pmatrix} u_1 + v_1 \\ v_1 \end{pmatrix} + \tilde{\beta} = \tilde{\Theta} \cdot \begin{pmatrix} u_2 + \frac{1}{\gamma}v_2 \\ \frac{1}{\gamma}v_2 \end{pmatrix} + \tilde{\beta}$$

Since,  $\succeq_2$  is regular and  $\tilde{\Theta}$  is non-negative, we conclude  $1 + b = \gamma$  for some  $b \geq 0$ . Hence,  $\gamma \geq 1$  as desired.

Suppose  $\succeq_2$  has more preference for commitment than  $\succeq_1$ , and  $\succeq_1$  and  $\succeq_2$  have the same self-control. Following the line of argument above, we obtain  $(u_2, v_2, \delta_2)$ , a representation of  $\succeq_2$  such that  $u_2 + v_2 = u_1 + v_1$  and  $v_2 = \gamma v_1$  for some  $\gamma > 0$ . Then, since  $\succeq_2$  has more preference for commitment than  $\succeq_1$ , (14) implies that there is a non-negative, non-singular  $\Theta$  and  $\beta$  such that

$$\begin{pmatrix} u_1 \\ v_1 \end{pmatrix} = \Theta \cdot \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} + \beta = \Theta \cdot \begin{pmatrix} u_1 + (1 - \gamma)v_1 \\ \gamma v_1 \end{pmatrix} + \beta$$

It follows from the regularity of  $\succeq_2$  that  $\gamma = \frac{1}{1-b}$  for some  $b \geq 0$  and since  $\gamma > 0$ , we have  $\gamma \geq 1$  as desired. *Q.E.D.*

## 12. Appendix D: Competitive Economies

**Proof of Proposition 1** We first show existence of equilibrium in a “truncated” economy in which from period  $\tau$  on every consumer is committed to the consumption of  $\omega_{it} \in C$  in every period. Let  $x_{it} = \{(\omega_{it}, x_{it+1})\}$  be the IHCP in which the agent is committed to consume the endowment. Let

$$x_{it}^\tau(p, b) = \begin{cases} \{(c, x_{it}(p, b')) \mid (c, b') \in B_t(p, b)\} & \text{if } t \leq \tau - 1 \\ \{(c_t, x_{it+1}) \mid (c, 0) \in B_t(p, b)\} & \text{if } t = \tau \\ x_{it} & \text{if } t \geq \tau + 1 \end{cases}$$

Let  $v_i^* : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be defined as

$$v_i^*(p_t, b) = \max_{\{c \mid p_t \cdot c \leq b\}} v_i(c)$$

and note that  $v_i^*$  is convex since  $v_i$  is convex. It is straightforward to show that the optimal consumption choices in periods  $1 \leq t \leq \tau$  given the IHCP  $x_1^\tau(p, \sum_{t=1}^\tau p_t \cdot \omega_{it})$  solve the

following maximization problem:

$$\begin{aligned} & \max_{\{c_t\}} \sum_{t=1}^{\tau} \delta^{t-1} \left[ u_i(c_t + v_i(c_t)) - v_i^* \left( p_t, \sum_{k=1}^{\tau} p_k \cdot \omega_{ik} - \sum_{k=1}^{t-1} p_k \cdot c_k \right) \right] \\ & \text{subject to} \\ & c_t \leq \sum_{k=1}^{\tau} p_k \cdot \omega_{ik} - \sum_{k=1}^{t-1} p_k \cdot c_k \end{aligned}$$

Note that the objective function is strictly concave, strictly increasing, and the feasible set of consumption choices is compact. Therefore, we may apply a standard argument (for example, Proposition 17.C.1 in Mas-Colell et.al.) to establish the existence of an equilibrium in the truncated economy.

Normalize equilibrium prices in the truncated economy so that  $p_{11} = 1$ . For periods  $t > \tau$  set  $p_{tl} = \delta^t$ . (Note that in the truncated economy, prices for  $t > \tau$  can be chosen arbitrarily as all consumers are committed to consuming their endowments.) Since endowments are bounded away from zero it follows that there is an  $h < \infty$  such that  $p_{tl} \leq \delta^{t-1}h$  for all  $(t, l)$  and every truncation  $\tau$ . This follows from a standard argument since  $v$  is increasing,  $u$  is continuous, strictly increasing and since aggregate endowment is bounded away from zero.

Let  $(p^\tau, \mathbf{c}^\tau)$  denote the equilibrium price and consumption pair of the  $\tau$  period truncation. Let  $(p, \mathbf{c})$  denote the limit of a convergent subsequence. We claim that  $(p, \mathbf{c})$  is an equilibrium for the economy. Since market clearing holds for every  $\tau$  it must hold also in the limit. It suffices therefore to show that  $\mathbf{c}_i$  solves the optimization problem for individual  $i$ . Observe that  $p^\tau \cdot \omega_i \leq h \sum_{t=1}^{\infty} \delta^{t-1} \sum_l \omega_{itl} < \infty$ . Hence, by the dominated convergence theorem,  $p^\tau \cdot \omega_i \rightarrow p \cdot \omega_i$  and therefore, the set of feasible consumption plans given the IHCP  $x_1^\tau(p, \sum_{t=1}^{\tau} p_t^\tau \cdot \omega_t)$  converge to the set of feasible consumption plan given  $x_1(p, \sum_{t=1}^{\infty} p_t \cdot \omega_t)$ . Now a routine argument (using the continuity of the agent's utility function) shows that  $\mathbf{c}_i$  is an optimal consumption choice from  $x_1(p, \sum_{t=1}^{\infty} p_t \cdot \omega_t)$ , that is,  $\mathbf{c}_i$  is optimal for  $i$  at prices  $p$ . *Q.E.D.*

**Example of Pareto Inefficiency** The utility function is

$$u(c_{t1}, c_{t2}) = \log c_{t1} - \lambda c_{t1} + c_{t2}$$

and

$$v(c_{t1}, c_{t2}) = \lambda c_{t1}$$

Let  $(\alpha, 1)$  be the endowment in every period and  $\lambda \in (0, \frac{1}{2})$ . Normalize prices so that the price of good 1 in period 1 is equal to 1. Straightforward calculations yield

$$p_{t1} = \delta^t$$

$$p_{t2} = \alpha \delta^t$$

The equilibrium welfare of the representative household as a function of  $\alpha$ , denoted  $W_\epsilon$  can be written as

$$W_\alpha = u(\alpha, 1) + v(\alpha, 1) - \max_{\{(a_1, a_2) \mid a_1 + \alpha a_2 \leq b\}} v(a_1, a_2) + \delta W_\alpha$$

$$W_\alpha = \log \alpha + 1 - \lambda b + \delta W_\alpha$$

where  $b = \frac{2\alpha}{1-\delta}$  is the equilibrium value of the agent's endowment. Hence,  $\frac{db}{d\alpha} = \frac{2}{(1-\delta)}$  and therefore  $\frac{dW_\alpha}{d\alpha} < 0$  at  $\alpha = 1$ . It follows that destroying  $\epsilon = 1 - \alpha$  units of good 1 each period increases welfare, for  $\epsilon$  sufficiently small.

**Proof of Proposition 2:** The necessary condition for an optimal plan is

$$u'(c_{it}) + \lambda_i v'(c_{it}) = \delta(1+r) (u'(c_{it}) + \lambda_i v'(c_{it}) - \lambda_i v'(\bar{b}_{it+1}))$$

where  $\bar{b}_{it+1}$  is the period  $t+1$  wealth of agent  $i$  in terms of period  $t+1$  consumption. In a steady state  $c_{it} = c_i$  and  $\bar{b}_{it} = \frac{c_i(1+r)}{r}$  yielding

$$u'(c_i) + \lambda_i v'(c_i) = \delta(1+r) \left( u'(c_i) + \lambda_i v'(c_i) - \lambda_i v' \left( \frac{c_i(1+r)}{r} \right) \right)$$

Choose  $i_0$  such that  $\lambda_{i_0} = \min \lambda_i$ . We define  $\bar{r}$  by the equation

$$\frac{\delta(1+\bar{r})-1}{\delta(1+\bar{r})\lambda_{i_0}} (u'(\bar{\omega}) + \lambda_{i_0} v'(\bar{\omega})) = v' \left( \frac{(1+\bar{r})\bar{\omega}}{\bar{r}} \right) \quad (18)$$

Note that  $\bar{r}$  is well-defined since the l.h.s. of the above equation is increasing in  $r$  and the r.h.s. is decreasing. Moreover, as  $r \rightarrow \infty$ , the l.h.s. converges to  $\frac{u'_i(\bar{\omega}) + \lambda_{i_0} v'_i(\bar{\omega})}{\lambda_{i_0}} > v'(\bar{\omega})$  and the r.h.s. converges to  $v'(\bar{\omega})$ . Hence, there is a unique  $\bar{r}$  that satisfies the above equation. Furthermore,  $\bar{r} > \frac{1-\delta}{\delta}$ .



Let  $r \in (\frac{1-\delta}{\delta}, \bar{r}]$ . For every  $r$  in that range there is a unique  $c_{i0}(r) \leq \bar{\omega}$  that satisfies (18) (i.e.,  $c_{i0}(r)$  such that (18) is satisfied when  $\bar{r}$  is replaced with  $r$  and  $\bar{\omega}$  is replaced with  $c_{i0}(r)$ ). To see this observe that since  $u'(c) + \lambda_{i_0} v'(c) \rightarrow \infty$  as  $c \rightarrow 0$ , the l.h.s. goes to infinity as  $c_{i0} \rightarrow 0$ . Since  $v$  is convex, the r.h.s. is bounded. For  $r \leq \bar{r}$ , we have

$$\begin{aligned} \frac{\delta(1+r) - 1}{\delta\lambda_i(1+r)} (u'(\bar{\omega}) + \lambda_i v'(\bar{\omega})) &\leq \frac{\delta(1+\bar{r}) - 1}{\delta\lambda_{i_0}(1+\bar{r})} (u'(\bar{\omega}) + \lambda_{i_0} v'(\bar{\omega})) \\ &= v' \left( \frac{(1+\bar{r})\bar{\omega}}{\bar{r}} \right) \leq v' \left( \frac{(1+r)\bar{\omega}}{r} \right) \end{aligned}$$

Since  $u' + \lambda_i v'$  is decreasing and  $v'$  is increasing, there is a unique solution,  $c_{i0}(r)$ .

Define  $c_i(r)$  for each  $i$  by replacing  $\lambda_{i_0}$  with  $\lambda_i$  in (18). Now, observe that  $c_i(r)$  is strictly increasing in  $r$ . To see this, note that holding consumption constant, the l.h.s. of (18) is strictly increasing in  $r$  while the convexity of  $v$  implies that the r.h.s. is decreasing. Then, the strict concavity of  $u + \lambda_i v$  ensures that consumption has to increase to maintain the equality of the two sides.

Thus,  $\sum_i c_i(r)$  is increasing in  $r$  with

$$\lim_{r \rightarrow \frac{1-\delta}{\delta}} \sum_i c_i(r) = 0 \text{ and } \sum_i c_i(\bar{r}) \geq \bar{\omega}$$

Continuity implies that there is a unique  $r^d$  such that  $\sum_i c_i(r^d) = \bar{\omega}$ .

Finally we need to show that  $c_i > c_j$  if  $\lambda_i < \lambda_j$ . Examine

$$\frac{\delta(1+r) - 1}{\delta\lambda_i(1+r)} (u'(c_i) + \lambda_i v'(c_i)) = v' \left( \frac{(1+r)c_i}{r} \right)$$

to see that if  $\lambda_i < \lambda_j$  and  $c_i \leq c_j$  then the necessary condition for an optimum must be violated for either  $i$  or  $j$ . *Q.E.D.*

### Stochastic Representative Agent Economy

As shown in the text, the equilibrium price of the productive asset is constant and satisfies

$$p = \delta(1 + p - 2\lambda p - \lambda p^2 - \lambda(\sigma^2 + 1)) \tag{19}$$

and hence

$$p = -\frac{1}{2\delta\lambda} \left( 1 - \delta + 2\delta\lambda - \sqrt{(1-\delta)^2 + 4\delta\lambda(1-\delta\lambda\sigma^2)} \right)$$

It can easily be checked that for  $\lambda < \frac{1}{1+\sigma^2}$  the price is strictly positive. Since  $\sigma^2 \leq 1$  this implies that for  $\lambda \in (0, 1/2)$  the price is strictly positive.

The risk free rate satisfies the following no-arbitrage condition:

$$\begin{aligned} u'(c_t) + v'(c_t) &= \delta(1+r)E[(u'(\tilde{c}_{t+1}) + v'(\tilde{c}_{t+1}))] \\ &\quad - \delta(1+r)E[v'(s_{t+1}(p(d) + d))] \end{aligned}$$

Using the equilibrium conditions  $c_t = D, \tilde{c}_{t+1} = d, s_{t+1} = 1$  this implies for our special case that

$$\begin{aligned} 1+r &= \frac{1}{\delta(1-\lambda(p+1))} \\ &= \frac{2}{1+\delta - \sqrt{(1-\delta)^2 + 4\delta\lambda(1-\delta\lambda\sigma^2)}} \end{aligned}$$

Equation (19) and  $p > 0$  imply that  $1 - \lambda(p+1) > 0$  and hence  $1+r > 0$ . To see that  $R-r$  is positive and increasing note that

$$R-r = \frac{1+p}{p} - \frac{1}{\delta(1-\lambda(p+1))} = \frac{(1+r)\delta(1+p)(1-\lambda(p+1)) - p}{p}$$

Equation (19) and some simplification yields

$$R-r = \lambda\delta\sigma^2 \frac{1+r}{p}$$

It remains to show that  $R-r$  is increasing in  $\lambda$ . Equation (19) implies

$$p(1-\delta) = \delta(1-2\lambda p - \lambda p^2 - \lambda(\sigma^2 + 1))$$

It follows that  $p$  is decreasing in  $\lambda$  and since  $p > 0$  that  $1-2\lambda p - \lambda > 0$ . We conclude that

$$\begin{aligned} \frac{\partial}{\partial \lambda} \left( \frac{1+r}{p} \right) &= \frac{\partial}{\partial \lambda} \left( \frac{1}{\delta p(1-\lambda(p+1))} \right) \\ &= \frac{(1-2\lambda p - \lambda)(-\partial p / \partial \lambda) + p + p^2}{\delta (p(1-\lambda(p+1)))^2} > 0 \end{aligned}$$

which in turn implies that  $R-r$  is increasing in  $\lambda$ .

*Q.E.D.*

## References

1. Bertsekas D. P. and S. E. Shreve (1978): *Stochastic Optimal Control: The Discrete Time Case*. New York: Academic Press.
2. Brandenburger, A. and E. Dekel (1993): “Hierarchies of Beliefs and Common Knowledge,” *Journal of Economic Theory*, 59, 189–199.
3. Brown A. and C. Percy (1995): *An Introduction to Analysis*. New York: Springer Verlag.
4. Bulow J. and K. Rogoff (1989): “Sovereign Debt: Is to Forgive to Forget?” *American Economic Review*, 79, 43–50.
5. Dekel, E., B. Lipman and A. Rustichini (2001): “A Unique Subjective State Space for Unforeseen Contingencies,” *Econometrica*, 69, 891–934.
6. Epstein, L. G. and J. A. Hynes (1983): “The Rate of Time Preference and Dynamic Economic Analysis.” *Journal of Political Economy*, 91: 611–635.
7. Epstein, L. G. and M. Peters (1999): “A Revelation Principle for Competing Mechanisms,” *Journal of Economic Theory*, 88, 119–160.
8. Epstein, L. G. and T. Wang (1996): “Beliefs about Beliefs without Probabilities,” *Econometrica*, 64, 1343–1374.
9. Epstein, L. G. and S. Zin, (1989): “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework” *Econometrica*, 57, 937–969.
10. Gul, F. and W. Pesendorfer, (2001): “Temptation and Self-Control,” *Econometrica*, 69, 1403–1436.
11. Kirby K. and R. J. Herrnstein (1995): “Preference Reversal Due to Myopia of Delayed Reward,” *Psychological Science*, 1995, 83–89.
12. Kocherlakota, N. R. (1996): “Reconsideration-Proofness: A Refinement for Infinite Horizon Time Inconsistency,” *Games and Economic Behavior*, 15, 33–54.
13. Kreps, D. and E. Porteus (1978): “Temporal Resolution of Uncertainty and Dynamic Choice Theory,” *Econometrica*, 46, 185–200.
14. Krusell P., and A. Smith (1999) “Consumption and Savings Decision with Quasi Geometric Discounting,” Unpublished Manuscript, University of Rochester.
15. Laibson, D. (1997): “Golden Eggs and Hyperbolic Discounting,” *Quarterly Journal of Economics*, 112, 443–477.
16. Lucas, R. E. “Asset Prices in an Exchange Economy,” *Econometrica*, 1978, 46(6): 1429–1445.

17. Mariotti, T. and M. Piccione (1999), “A Hierarchical Construction of a Compact Possibility Model,” Unpublished Manuscript, London School of Economics.
18. Mertens, J-F. and S. Zamir (1985): “Formulation of Bayesian Analysis for Games with Incomplete Information,” *International Journal of Game Theory*, 14, 1–29.
19. O’Donoghue, T. and M. Rabin (1999): “Doing It Now or Later,” *American Economic Review*, 89, 103–125.
20. Parthasarathy K. (1970): *Probability Measures on Metric Spaces*. New York: Academic Press.
21. Phelps, E. S. and R. A. Pollak (1968): “On Second-Best National Saving and Game-Equilibrium Growth”, *Review of Economic Studies*, 35, 185–199.
22. Piccione, M. and A. Rubinstein (1997): “On the Interpretation of Decision Problems with Imperfect Recall,” *Games and Economic Behavior*, 20, 3–24.
23. Rabin M. (1997): “Psychology and Economics,” Unpublished Manuscript, Department of Economics, University of California.
24. Strotz, R. H. (1955): “Myopia and Inconsistency in Dynamic Utility Maximization,” *Review of Economic Studies*, 23, 165–180.

## Footnotes

<sup>1</sup> We thank Pierre-Olivier Gourinchas, Per Krusell and Robert Shimer for their suggestions and advice. We thank four anonymous referees and a co-editor for their comments. This research was supported by grants from the National Science Foundation and by the Sloan Foundation.

<sup>2</sup> See, for example, Strotz (1955), Laibson (1997), O'Donoghue and Rabin (1998), Krusell and Smith (1999)

<sup>3</sup> See Parthasarathy (1970).

<sup>4</sup> See Brown and Pearcy (1995) p. 222.

<sup>5</sup> This observation is due to Kreps (1978) who provides the argument for the finite setting. In our setting, note that Axiom 2 ensures that the restriction of  $\succeq$  to singleton sets is a continuous preference relation on  $\Delta$ . Hence there is a continuous utility function  $u : \Delta \rightarrow \mathbb{R}$  representing this relation. For finite  $x, y$  we conclude that  $x \succeq y$  iff  $\max_x u(\mu) \geq \max_y u(\mu)$  using Kreps' argument. To extend the representation to arbitrary compact sets, let  $\mu$  be a maximizer of  $u$  in the compact set  $x$ . Note that for finite subsets  $y, y'$  of  $x$  containing  $\mu$  we have  $y \sim y'$ . Continuity (in the Hausdorff topology) then implies  $x \sim y$ .

<sup>6</sup> The domain of preferences used here is closely related to the domain of preferences used in Kreps and Porteus (1978). Their "descriptive approach" defines preferences on the finite horizon analogue of  $\Delta$ , that is, lotteries over current consumption and continuation problems for the next period. They use this framework to analyze agents that may have a preference for early or late resolution of uncertainty.

<sup>7</sup> To see this, note that  $x \succ y$  implies  $\{(c, x)\} \succ \{(c, y)\}$  by Axiom 6. Axiom 4 then implies  $\{\alpha(c, x) + (1 - \alpha)(c, z)\} \succ \{\alpha(c, y) + (1 - \alpha)(c, z)\}$  and by Axiom 7  $\{(c, \alpha x + (1 - \alpha)z)\} \succ \{(c, \alpha y + (1 - \alpha)z)\}$ . Applying Axiom 6 again then yields the desired conclusion. Note that Axiom 4\* applied to singleton sets yields Axiom 4.

<sup>8</sup> This axiom was first used by Dekel, Lipman and Rustichini (2001) in their analysis of preferences over sets of lotteries.

<sup>9</sup> Conversely, if  $\succeq$  is represented by  $(u, v, \delta)$  with  $u$  and  $v$  not constant then  $\succeq$  is a DSC preference with a preference for commitment at some  $z$ .

<sup>10</sup> There is a close connection between our dynamically consistent interpretation and Machina's (1989) dynamically consistent interpretation of non-expected utility preferences. The "changing tastes" interpretation is analogous to the consequentialist view in Machina (1989). The changing tastes interpretation views period  $t > 0$  preferences as situation-independent. The consequentialists in Machina (1989) view preferences to be history-independent. In the case of non-expected utility, the two views imply different behavior. In our case, the distinction is more subtle. While period  $t > 0$  behavior is the same, the two approaches differ when the welfare implications of policies are evaluated.

<sup>11</sup> Recall that  $(c, x)$  denotes a *lottery* that yields the consumption  $c$  and the continuation problem  $x$  with probability 1.

<sup>12</sup> As an alternative, O'Donoghue and Rabin (1999) advocate "naive" behavior as a solution concept to this game. Naive agents incorrectly predict future behavior to maximize current period preferences.

<sup>13</sup> The argument that re-negotiation is particularly plausible in time-inconsistent decision problems is due to Kotcherlakota (1996).

<sup>14</sup> For a related critique of the use of Nash equilibrium when modeling a (different) departure from fully rational behavior, see Piccione and Rubinstein (1997)