A NOTE ON PAYMENTS IN EXPERIMENTS OF INFINITELY REPEATED GAMES WITH DISCOUNTING

ARUN G. CHANDRASEKHAR[†] AND JUAN PABLO XANDRI[⋆]

(PRELIMINARY AND INCOMPLETE)

Abstract. It is common for researchers studying repeated and dynamic games in a lab experiment to pay participants for all rounds or a randomly chosen round. We argue that these payment schemes typically implement different set of subgame perfect equilibria (SPE) outcomes than the target game. Specifically, paying a participant for all rounds or for a randomly chosen round makes the game such that early rounds matter more to the agent, by lowering discounted future payments. In addition, we characterize the mechanics of the problems induced by these payment methods. We are able to measure the extent and shape of the distortions. We also establish that a simple payment scheme, paying participants for the last (randomly occurring) round, implements the game. The result holds for any dynamic game with time separable utility and discounting. A partial converse holds: any payment scheme implementing the SPE should generically be history and time independent and only depend on the contemporaneous decision.

JEL Classification Codes: C90, C91, C92

Keywords: Payment in experiments, Experimental economics

1. Introduction

Many lab experiments study behavior in repeated and dynamic games. The goal is to study behavior in a game specified by theory, which we call the target game. Typically, a participant plays a round of a game which then continues to the next round with a given probability. In order to incentivize participant behavior, the experimenter pays the participant as a function of the history of play. Initially, the experiments in the literature paid participants for every round that they took part in. However, in order to compensate for self-insurance across rounds, some experiments moved to paying individuals for one randomly chosen round.[1] Examples of studies in the literature that either paid participants for all rounds or random rounds include Murnighan and Roth (1983), Palfrey and Rosenthal

Date: December 2011.

We thank Abhijit Banerjee, Juan Dubra, Matthew O. Jackson, Anil Jain, Juan Passadore, Rob Townsend, Ivan Werning, and Luis Zermeño for helpful discussions. Chandrasekhar is grateful for support from the National Science Foundation GRFP..

[†]Massachusetts Institute of Technology. Email: `arunc@mit.edu`.

[†]Massachusetts Institute of Technology. Email: `jxandri@mit.edu`.

[1]The argument offered was that "this payment structure prevents individuals from self-insuring income risk across rounds and that the utility maximization problem of the experiment matches that of the theoretical model" (Fischer, 2010).

(1994), Aoyagi and Fréchette (2009), Bo (2005), Engle-Warnick and Slonim (2004, 2006a,b), Charness and Genicot (2009), Duffy and Ochs (2009), Fischer (2011), Chandrasekhar et al. (2011).[2]

In this paper, we present a general framework to think about payment in experiments. We argue that the payment scheme used may induce a game different from the target game; researchers must be cognizant of the game they induce their participants to play. As such, researchers ought to check whether their mode of payment implements the same subgame perfect equilibria (SPE) outcomes as the target game, and if it does not, they ought to account for this in their analysis.

Our core results are as follows. First, we establish a test based on Abreu (1988) to see if a payment scheme implements a particular SPE: the set of implementable actions at any given history must be history independent. Second, we argue that there is a simple payment scheme, paying individuals for the last round in which they participated, that implements the target game. Third, we discuss the payment scheme which pays individuals for all rounds, show that it does not implement the target game, and even induces individuals to become asymptotically indifferent about their actions. Fourth, we discuss the payment scheme which pays individuals for a round chosen at random and argue that it does not alleviate the problems. We show that it has a set of implementable outcomes that is generally different from the target game; people will typically behave differently along the equilibrium path. Moreover, we characterize the mechanics of the behavior induced by this payment scheme. Paying individuals for a randomly chosen round induces them to discount the future too much, from any period, relative to the target game. Individuals also become asymptotically indifferent between decisions.

To make this more concrete, we provide examples where the equilibrium is changed drastically simply when the researcher changes the payment scheme. Consider repeated Prisoner's Dilemma in which, for a certain level of discounting, under the target game cooperation is sustained in equilibrium. Consequently, paying participants for the last round only generates exactly the same SPE. However, paying for a randomly chosen round rules out cooperation in equilibrium for a certain range of the discounting parameter.

In general, our results demonstrate that distortions are sizeable. For any discount rate and any dynamic game, the "virtual" net present value (NPV) of a constant stream of consumption at the first round, under round at random payment, would be at the most half of the actual NPV. In particular, if the discount rate is 0.95, the "virtual" NPV would be 30% less than the value under the theoretical model. Moreover, paying in all rounds

[2]Charness and Genicot (2009) note that paying individuals for a randomly chosen game introduces a different form of discounting and show that the discount factor converges to the right one mandated by the target game. However, they argue that this alone suffices to show that payment for a randomly chosen round approximately implements the target game – a claim which we show to be false.

creates distortions when we allow for curvature in the utility for wealth.[3] In an example, we show that for utility arbitrarily close to linear (CES utility with almost zero elasticity of intertemporal substitution) the implementable outcomes are indeed very different and have very different asymptotic properties. Finally, we note that the implementation result can be easily generalized for any dynamic game with discounting.

A windfall of our results is that it allows a researcher to correct her theoretical predictions if she was interested in estimating a structural model form the lab experiment data. By explicitly characterizing the discounting induced by the payment mechanisms as well as its impact on the induced game, we suggest that a researcher interested in structurally estimating parameters simply use the corrected model when performing structural estimation.

In Section 2 we present a simple example of repeated Prisoner's Dilemma to provide intuition about our results. In this example we show an example where the round at random payment scheme induces a game which generates potentially different equilibrium paths than what would have transpired under the target game. Specifically, if individuals were payed correctly, cooperation would have been sustained in equilibrium but for a range of the parameter space, employing the wrong payment scheme rules out cooperation. Section 3 establishes the general framework and develops a test of whether a payment scheme implements the target game. Section 4 demonstrates that the last period payment scheme implements the target game. In Section 5 we study the payment schemes from the previous literature, payment for all rounds and payment for a round chosen at random, and characterize the mechanics of the problem. Section 6 discusses our results in the context of a simple model of savings behavior. Section 7 concludes.

2. A Simple Example

We begin with a simple example of repeated Prisoner's Dilemma. Let the stage game be described by the following per-period payoffs in utility space, where $a > 1$, $a - b < 2$, as in Kreps et al. (1982).

	C	D
C	1,1	$-b, a$
D	$a, -b$	0,0

It is easy to see that under grim trigger punishment, cooperation is enforceable if and only if $\beta > 1 - \frac{1}{a}$.[4] Let us suppose that $a = 3$, $b = 3/2$, and $\beta = 7/10$. Cooperation is sustainable if and only if $\beta \geq \frac{2}{3}$.

Assume that a participant is told that she must play the game in period 1, and that the game will continue with probability $\frac{2}{3}$ (in every period when the game is still not over); she

[3]Under risk-neutrality, all round payment would implement the same set of SPE outcomes, though we show in Section 6.2 that this result is not robust.

[4]This follows from $(1-\beta) + \beta > (1-\beta)\, a \iff \beta > 1 - \frac{1}{a}$

will continue to play until the game ends. She will be paid her payoff from a (uniformly) randomly selected round.

To show that the equilibria of the theoretical game, and the experimental game are not the same, suppose that cooperating forever, and playing grim trigger in any history different from $h^t = (C, C)^t$, is an equilibrium. This will generate a contradiction.

In round 10, after having cooperated for 9 periods, if players have gotten there, the payoff of playing to the proposed equilibrium is 1. This is because the history will have been $(C, C)^t$, which paid 1 for each player. However, the payoff of playing D is larger than 1.

Finish at t	10	11	12
Probability	$1-\beta$	$(1-\beta)\beta$	$(1-\beta)\beta^2$
Payment	$\frac{9}{10} + \frac{1}{10} \cdot 3$	$\frac{9}{11} + \frac{1}{11} \cdot 3 + \frac{1}{11} \cdot 0$	$\frac{9}{12} + \frac{1}{12} \cdot 3 + \frac{1}{12} \cdot 0$

One can check that

$$\sum_{t=0}^{\infty} (1-\beta)\,\beta^t \left(\frac{9}{10+t} + \frac{3}{10+t}\right) = \sum_{t=0}^{\infty} \left(\frac{7}{10}\right)^t \cdot \frac{3}{10} \cdot \left(\frac{9}{10+t} + \frac{3}{10+t}\right) = 1.0116$$

so cooperation is not an equilibrium.

We repeat the exercise for several a, b which define a Prisoner's Dilemma. We present threshold values of β above which cooperation can be sustained.

	Target Threshold β	Implied Threshold β
$a = 2$	0.50	0.57
$a = 3$	0.67	0.82
$a = 4$	0.75	0.90
$a = 5$	0.80	0.92
$a = 6$	0.83	0.94

The exercise shows that the round at random payment scheme induces a game which generates potentially different equilibrium paths than what would transpire under the target game. This will be important to keep in mind, given certain convergence properties we will later show. The incentive compatibility constraint faced by an individual under the random round payment scheme converges to that of the target game (as noted in Charness and Genicot, 2009). However, we have shown, this does not imply that the same equilibrium, nor equilibrium path, is implemented in the induced game. That is, simple convergence of the incentive compatibility is not sufficient to generate convergent equilibrium behavior. A more detailed discussion of this phenomenon is provided in Section 5.2.

3. Framework

3.1. **Setup and Notation.** The researcher is interested in conducting an experiment to test behavior in a repeated game Γ, specified as

$$\Gamma := \left\{ \left\{ A_i,\ u_i : \mathbb{R} \to \mathbb{R},\ r_i : \prod_{i=1}^{n} A_i \to \mathbb{R} \right\}_{i=1}^{n},\ \beta \in (0,1) \right\}$$

where A_i is the strategy space for agent i, $a_t = \left(a_t^1, a_t^2, ..., a_t^n\right) \in \prod_{i=1}^n A_i$ is an action profile at time t, and β is a common discount rate for all agents. The overall utility of agent i over sequences $\mathbf{a} = \{a_t\}_{t=0}^{\infty} \in A^{\infty}$ is given by the time separable utility,

$$U_i\left(\mathbf{a}\right) = (1-\beta)\sum_{t=1}^{\infty}\beta^t u_i\left(r_i(a_t)\right)$$

Here $r_i(a_t)$ is the monetary reward of the game and the function $u_i\left(c\right)$ is the utility of wealth in the theoretical model that the research wants to test. This environment also includes models in which there is randomness in every period, in which case we can think of the utility functions u_i as designating expected utility functions (a la Von Neumann-Morgenstern).

Let $\mathcal{H} = \cup_{t=1}^{\infty} A^t$ be the set of all possible histories and define a *pure strategy for agent* i as a function $\sigma_i : \mathcal{H} \to A_i$, which specifies after each possible history of play $h^{t-1} = (a_1, a_2, ..., a_{t-1})$, an action $a_{i,t} = \sigma_i\left(h^{t-1}\right)$. We will also denote $\sigma_i|h^t$ to be the strategy profile conditional on history h^t. Given a strategy profile σ, the *outcome of* σ, henceforth denoted by $\mathbf{a}\left(\sigma\right) \in A^{\infty}$, is the sequence of actions prescribed by the strategy profile. That is, $\mathbf{a}\left(\sigma\right) = \left(\sigma_1\left(\varnothing\right), \sigma_2\left(\sigma_1\left(\varnothing\right)\right), ...\right) := \{a_t\left(\sigma\right)\}_{t=1}^{\infty}$.

A strategy profile $\sigma = (\sigma_1, \sigma_2, ..., \sigma_n)$ is a *Subgame Perfect Equilibrium of* Γ, and we write $\sigma \in \mathbf{SPE}\left(\Gamma\right)$, if and only if for all histories $h^t \in \mathcal{H}$, all agents $i = 1, 2, ..., n$ and all alternative strategies σ_i', we have that $U_i\left(\mathbf{a}\left(\sigma \mid h^t\right)\right) \geq U_i\left(\mathbf{a}\left(\sigma_i', \sigma_{-i} \mid h^t\right)\right)$. That is, a strategy profile is a subgame perfect equilibrium of Γ if it is a Nash equilibrium for all subsequent subgames. We say that sequence $\mathbf{a} = \{a_t\}_{t=0}^{\infty}$ is an *implementable outcome* if it is the outcome of some SPE profile σ; i.e. $\exists \sigma \in \mathbf{SPE}\left(\Gamma\right) : \mathbf{a} = \mathbf{a}\left(\sigma\right)$. In general, we will define the *set of implementable outcomes of a game* Γ as

$$\mathbf{O}\left(\Gamma\right) := \bigcup_{\sigma \in \mathbf{SPE}(\Gamma)} \mathbf{a}\left(\sigma\right) \subseteq A^{\infty}.$$

Given a history of play h^{t-1}, we say that an *action profile* $a_t \in A$ *is implementable at* h^{t-1} if and only if there exist some implementable outcome $\mathbf{a}$ such that $\mathbf{a}_t = a_t$. We will also denote the set of all implementable action profiles after h^{t-1} as $O_t\left(\Gamma, h^{t-1}\right) \subseteq A$. If we add the condition that the equilibrium path of play from $t+1$ is $\mathbf{a}' = (a_{t+1}, a_{t+2}, ...) \in \mathbf{O}\left(\Gamma\right)$, then the set of implementable outcomes at time t is reduced, and we write $a_t \in O_t\left(\Gamma, h^{t-1}|\mathbf{a}'\right) \subset O_t\left(\Gamma, h^{t-1}\right)$.

Finally, let $V\left(\Gamma\right)$ be the set of SPE equilibrium values,

$$V\left(\Gamma\right) := \left\{v \in \mathbb{R}^n \ : \ v_i = U_i\left(\mathbf{a}\right) \text{ where } \mathbf{a} \in \mathbf{O}\left(\Gamma\right)\right\}.$$

3.2. **Payment Schemes and a Test of Implementation.** The researcher wants to test Γ, which has an infinite horizon. As such, she has to design an alternative finite repeated game $\widehat{\Gamma}$ to test the predictions of game Γ. In particular, it is a well known fact that the infinite horizon nature of Γ can be replicated by a game that ends in T periods, where T is a

geometric random variable with parameter β, whose realization is not known to the players, so that $\mathrm{P}(T=t)=(1-\beta)\beta^t$. We write $T \sim G(\beta)$.

In order to test the predictions of game Γ, the researcher has to choose a *payment scheme* $\mathbf{R}=\{R_{i,t}(h^t)\}_{i,t}$ that specifies, for each history h^t, the payment that agent i should receive, in terms of a distribution $R_{i,t}(h^t)$ of monetary rewards *at the end of the game*. For example, she might choose a payment scheme that runs the stage game for T periods, where $T \sim G(\beta)$ and such that pays only one of the previous monetary rewards with equal probability, once the game is terminated: i.e.

$$R_{i,t}(h^t)=\begin{cases} r_i(a_s) \text{ with prob. } \frac{1}{t} \text{ for all } s \leq t & \text{if} \quad t=T \\ 0 & \text{otherwise.} \end{cases}$$

We will call this payment scheme *round at random* and we will denote it $\boldsymbol{R}_{RAR}$. Once the payment scheme is chosen, it generates a new induced repeated game, which we will call $\widehat{\Gamma}(\boldsymbol{R})$, which has the same strategy space as Γ but evaluates action sequences differently. At any history h^t, the individual evaluates future sequences according to:

$$\widehat{U}_i\left(\mathbf{a} \mid h^t\right)=E\left\{u_i\left[R_{i,T}\left(h^T \mid h^t\right)\right]\right\}$$

If the researcher wants to test the predictions of game Γ, the payment scheme used *must not change the set of subgame perfect equilibria of the original game* Γ. That idea is formalized in the next definition.

Definition 3.1 (Implementation)**.** We say a *payment scheme* $\boldsymbol{R}=\{R_{i,t}(h^t)\}_{i,t}$ *implements* Γ if the set of implementable outcomes from Γ and from $\widehat{\Gamma}(\boldsymbol{R})$ coincide, i.e. $\mathbf{O}\left(\widehat{\Gamma}(\boldsymbol{R})\right)=\mathbf{O}(\Gamma)$.

We refer to such schemes as *implementing schemes*. Given a payment scheme $\boldsymbol{R}$, we are interested in studying whether it implements Γ. We use a powerful implication of Abreu (1988) about simple profiles to establish a test of whether $\boldsymbol{R}$ implements Γ.

Assumption 1. $\mathbf{SPE}(\Gamma) \neq \varnothing$.

Assumption 2. $A=\prod_{i=1}^{i=n} A^i$ *is a compact set.*

Assumption 3. $u_i(r_i(\cdot))$ *are continuous functions of* $a \in A$ *for all* $i=1,2,...,n$.

Proposition 3.1. *Under Assumptions 1, 2, 3, if* $\boldsymbol{R}$ *implements* Γ*, then the set of implementable action profiles at history* h^t *is time and history independent; i.e.* $O_t\left(\widehat{\Gamma}(\boldsymbol{R}), h^{t-1}\right)$ *is not a function of* h^{t-1} *nor a function of* t.

This proposition builds on one of the key insights in Abreu (1988). Take a candidate SPE outcome $\mathbf{a}$ in game Γ such that from tomorrow on, if we continue to be on the equilibrium path, it prescribes outcome $\bar{\mathbf{a}}'$ from $t+1$ on. If an individual deviates, we can always prescribe the optimal punishment, which is history and time independent. Contemporaneous utility

in a repeated game is also history and time independent. It follows that the set of action profiles which are implementable at time t must also be history independent. Because $\Gamma$and $\widehat{\Gamma}(\boldsymbol{R})$ share the same set of outcomes, the result obtains.

By making payments $R_{i,t}(h^t)$ depend on t and on past history h^{t-1} we will typically violate this property. For instance, we show that $\boldsymbol{R}_{RAR}$ violates this property since payments at any history h^t depend on the previous actions h^{t-1} and also on t, which governs the distribution of the rewards.

4. Last Round Payment

We first establish that for any given game Γ we can find a payment scheme that not only implements it, but also has exactly the same set of subgame perfect equilibria. Define the scheme $\boldsymbol{R}_{last}$ by

$$R_{i,t}(h^t) := \begin{cases} r_i(a_t) \text{ with prob. } 1 & \text{if} \quad t = T \\ 0 & \quad\quad \text{otherwise,} \end{cases}$$

where $T \sim G(\beta)$. Here we pay agents for only the last period of the game. More specifically, having arrived at a history h^{t-1} at time t, the payment scheme prescribes that with probability $1-\beta$ the game ends and agents get as final reward $r_i(a_t)$. With probability β the game continues at least for one more period, and whatever was played at time t does not enter into the final payment. This payment scheme does not depend on the past history h^{t-1}, so that the implementability test implied by Proposition 3.1 is satisfied. The next proposition presents the main result of the paper.

Proposition 4.1. $\mathbf{SPE}(\Gamma) = \mathbf{SPE}\left(\widehat{\Gamma}(\boldsymbol{R}_{last})\right)$ *which implies that* $\boldsymbol{R}_{last}$ *implements* Γ. *Moreover,* $V(\Gamma) = V\left(\widehat{\Gamma}(\boldsymbol{R}_{last})\right)$.

Proof. The proof follows from the fact that, for any history h^t and any sequence $\mathbf{a} \mid h^t$ we have that

$$\widehat{U}_i\left(\mathbf{a} \mid h^t\right) = E_T\left\{u_i(r_i(a_{t+T}))\right\} = \sum_{s=0}^{\infty}(1-\beta)\,\beta^t u_i(r_i(a_{t+s})) = U_i\left(\mathbf{a} \mid h^t\right)$$

because T, being geometric, has no memory. Since in both games Γ and $\widehat{\Gamma}(\boldsymbol{R}_{last})$ utilities are measured identically for all subgames starting at any history h^t, it follows that $\mathbf{SPE}(\Gamma) = \mathbf{SPE}\left(\widehat{\Gamma}(\boldsymbol{R}_{last})\right)$ and $V(\Gamma) = V\left(\widehat{\Gamma}(\boldsymbol{R}_{last})\right)$. □

This argument is easily generalizable. It is easy to see that $\boldsymbol{R}_{last}$ will implement *any* multistage game with observable actions with time separable utility and common discount factors.[5]

[5]The utility functions, strategy spaces, and common discount factors can all be history dependent. For instance, this payment scheme is implementing for dynamic games with capital accumulation, savings, etc., which are often of interest.

5. Payment Schemes in the Literature

We now turn to payment schemes used in the literature. A number of experiments have paid participants for their total earnings across all rounds in the game. Noting that the curvature of utility of wealth may cause the participant to deviate from the modeled behavior, experiments have turned to paying participants for a single randomly chosen round. In this section we first review the payment for all rounds scheme and explicitly characterize the curvature effect. In doing so we develop tools that allow us to turn to the round at random payment scheme and analyze explicitly how it alters participant behavior.

5.1. **Payment for All Rounds.** In this section we study the payment scheme wherein the agent is paid for every round of participation. Paying individuals for all rounds may not implement the model Γ.[6] We establish that the payment schemes may significantly weaken the incentives of the participants as the number of rounds played increases.

The payment scheme $\boldsymbol{R}_{all}$ is given by

$$R_{i,t}(h^t) = \begin{cases} \sum_{s=0}^{s=t} r_i(a_s) \text{ with prob. } 1 & \text{if} \quad t = T \\ 0 & \quad\quad \text{otherwise} \end{cases}$$

where $T \sim G(\beta)$. Given an outcome $\mathbf{a} = \{a_t\}_{t=0}^{\infty}$ we define $R_{i,t}(\mathbf{a})$ as the accumulated rewards up to time t for agent i; i.e. $R_{it}(\mathbf{a}) = \sum_{s=0}^{t} r_i(a_s)$ and let $R_t(\mathbf{a}) = (R_{1,t}(\mathbf{a}), R_{2,t}(\mathbf{a}), ..., R_{n,t}(\mathbf{a}))$ the vector of all accumulated rewards up to time t. The following proposition gives a characterization of the set of implementable outcomes of $\widehat{\Gamma}(\boldsymbol{R}_{all})$.

Proposition 5.1. *Under Assumptions 1, 2, 3, a sequence* $\mathbf{a} = \{a_t\}_{t=0}^{\infty}$ *is implementable if and only if for all* i *and all* t*:*

$$a_{i,t} \in \underset{\tilde{a}_i \in A_i}{\operatorname{argmax}} (1-\beta) u_i [R_{i,t-1}(\mathbf{a}) + r_i(\tilde{a}_i, a_{-i,t})] + \beta W_i(\tilde{a}_i, a_{-i,t}) \tag{5.1}$$

where

$$W_i(\tilde{a}_i, a_{-i,t}) = \begin{cases} \underline{W}_i(R_t(\mathbf{a})) & \text{if} \quad \tilde{a}_i \neq a_{i,t} \\ (1-\beta)\sum_{s=0}^{\infty} \beta^s u_i(R_{i,t+s}(\mathbf{a})) & \quad\quad \text{otherwise} \end{cases}$$

where the function $\underline{W}_i(R_t)$ *is the value of the optimal punishment for agent* i *starting from history* $h^t = (a_1, a_2, ..., a_t)$*.*

Proposition 5.1 characterizes the distortion due to $\boldsymbol{R}_{all}$. Since agents receive the payment for the experiment only when the experiment ends, the amount earned up to time t generates a stock of earnings not yet consumed. If there is some curvature in the utility function u_i, then the stock of unconsumed earnings may affect incentives of agent i in all subsequent rounds. In particular, if utility over monetary rewards were concave, we should expect to see a diminishing marginal utility of wealth as t increases, which would weaken incentives in

[6]This has been noted in Charness and Genicot (2009) and Fischer (2010), among others.

the long run. We mention that if utility was linear in earnings, this payment scheme would not cause problems as past earnings would drop from condition (5.1).

We formalize these intuitions in Proposition 5.2. Let

$$r_i^{\max} = \max_{a \in A} r_i(a) \text{ and } r_i^{\min} = \min_{a \in A} r_i(a)$$

be the best and worst possible stage rewards for agent i, and suppose that $r_i^{\min} \geq 0$.

Proposition 5.2. *Suppose that $u^i(\cdot)$ is a concave differentiable function and $r_i^{\min} \geq 0$.*

(1) *If $u_i' \geq 0$ and $u_i'' \leq 0$ for agent i, then the range of values for contemporaneous and continuation utilities is decreasing over time. Specifically, for any history h^t and any pair of continuation sequences $\mathbf{a}, \mathbf{a}' \in A^\infty$ we have that*

$$\left|\hat{U}_i\left(\mathbf{a} \mid h^t\right) - \hat{U}_i\left(\mathbf{a}' \mid h^t\right)\right| \leq u_i'\left[(t+1)\, r_i^{\min}\right]\left(r_i^{\max} - r_i^{\min}\right).$$

(2) *If u_i satisfies the Inada condition $u_i'(\infty) = 0$ and $r_i^{\min} > 0$, then as $t \to \infty$*

$$\sup_{\mathbf{a},\mathbf{a}' \in A^\infty} \left|U_i\left(\mathbf{a} \mid h^t\right) - U_i\left(\mathbf{a}' \mid h^t\right)\right| \to 0.$$

(3) *If utilities are affine for all agents (i.e. $u_i(c) = \kappa_i + \gamma_i c$) then $\boldsymbol{R}_{all}$ implements Γ.*

Proposition 5.2 illustrates the nature of the distortion caused by $\boldsymbol{R}_{all}$. As time passes the amount by which an agent's utility changes must be decreasing. The payment scheme is only implementing if the participants are modeled as risk-neutral.

5.2. **Payment for a Randomly Chosen Round.** In this subsection we argue that the round at random payment scheme introduced in Section 3.2 also fails the test of Proposition 3.1 and does not implement Γ. Moreover, we characterize the behavior induced by this payment scheme. First, it has a set of implementable outcomes that is generally different from the target game. Participants typically behave differently along the equilibrium path. Second, individuals discount the future too much in any period. Third, individuals to become asymptotically indifferent between their choices.

Recall the functions

$$\eta(\beta, t) = \sum_{k=0}^{\infty} \frac{\beta^k}{t+k} \text{ and } \eta_\beta(\beta, t) = \frac{\partial \eta(\beta,t)}{\partial \beta} = \sum_{k=0}^{\infty} \frac{k\beta^{k-1}}{t+k} \tag{5.2}$$

described in section 2. We catalog their properties in Appendix B. We show in Lemma B.3 that utility of outcome $\mathbf{a}$ from time $t = 0$ onward can be written as

$$\widehat{U}_i(\mathbf{a}) = (1-\beta) \sum_{t=0}^{\infty} \eta(\beta, t)\, \beta^t u_i\left[r_i(a_t)\right] \tag{5.3}$$

and at any history h^t

$$\widehat{U}_i\left(\mathbf{a} \mid h^t\right) = (1-\beta)\,\eta(\beta,t)\sum_{k=0}^{t} u_i\left[r_i\left(a_k\right)\right] + \beta\left\{(1-\beta)\sum_{s=0}^{\infty}\beta^s\eta(\beta,t+1+s)\,u_i\left[r_i\left(a_{t+s+1}\right)\right]\right\}. \tag{5.4}$$

The key change that we see in (5.3) is that the implied discount factor is $\eta(\beta,t)\,\beta^t$ instead of just β^t. Individuals in game $\widehat{\Gamma}(\boldsymbol{R}_{RAR})$ discount future flows of utility too rapidly in the induced game relative to Γ. We can see this from the fact that $\lim_{t\to\infty}\eta(\beta,t) = 0$ (see Lemma B.1) and therefore

$$\lim_{t\to\infty}\frac{\eta(\beta,t)\,\beta^t}{\beta^t} = 0.$$

As a practical matter, we should expect agents to behave much more impatiently than the actual model would predict. This comes from the fact that, when choosing a_1 at $t=1$, the agent should internalize not only the fact that she should receive $(1-\beta)\,u_i\left[r_i\left(a_1\right)\right]$ utils at $t=1$, but that this also affects the expected utility at time $t=2$ by $(1-\beta)\,\beta\frac{1}{2}u_i\left[r_i\left(a_1\right)\right]$, at time $t=3$ by $(1-\beta)\,\beta^2\frac{1}{3}u_i\left[r_i\left(a_1\right)\right]$, and so on. Ultimately, this increases the weight of time $t=1$'s decision on lifetime utility as it shows up in every subsequent utility computation. The next result illustrates the shape of the distortions caused by round at random payment.

Proposition 5.3 (Implementable outcomes in $\widehat{\Gamma}(\boldsymbol{R}_{RAR})$)**.** *Suppose assumptions 1, 2 and 3 hold. Under $\boldsymbol{R}_{RAR}$ a sequence $\{a_t\}_{t=0}^{\infty}$ is an implementable outcome of $\widehat{\Gamma}(\boldsymbol{R}_{RAR})$ if and only if for all $t\in\mathbb{N}$ and all $i\in\{1,2,...,n\}$,*

$$a_{i,t} \in \underset{\tilde{a}_i\in A_i}{\operatorname{argmax}}\,(1-\beta)\,u_i\left[r_i\left(\tilde{a}_i,a_{-i,t}\right)\right] + \beta W_i\left(\tilde{a}_i,a_{-i,t}\right) \tag{5.5}$$

with

$$W_i\left(\tilde{a}_i,a_{-i,t}\right) = \begin{cases} \underline{W}^i(t) & \text{if } \tilde{a}_i \neq a_{i,t} \\ (1-\beta)\sum_{s=0}^{\infty}\left[\frac{\eta(\beta,t+1+s)}{\eta(\beta,t)}\right]\beta^s u_i\left[r_i\left(a_{t+s}\right)\right] & \text{otherwise} \end{cases},$$

where $\underline{W}_i(t)$ is the value of agent i's optimal punishment in game Γ_t (as defined in Lemma B.4),

$$\underline{W}_i(t) = \min\left\{v^i : \exists v^{-i}\in\mathbb{R}^{n-1} \text{ with } \left(v^i,v^{-i}\right)\in V\left(\Gamma_t\right)\right\}.$$

Proposition 5.3 illustrates the nature of the distortion over the set of implementable outcomes generated by $\boldsymbol{R}_{RAR}$. Condition (5.5) is nearly identical to the implementability condition of game Γ – the crucial distinction being that future rewards are further discounted by the term $\eta(\beta,t)^{-1}\cdot\eta(\beta,t+s+1)$. As η is decreasing in t, we have that $\eta(\beta,t)^{-1}\cdot\eta(\beta,t+s+1)<1$. This immediately implies that for all s,

$$\eta(\beta,t)^{-1}\cdot\eta(\beta,t+s+1)\,\beta^s < \beta^s.$$

Consequently a participant is more impatient in $\widehat{\Gamma}(\boldsymbol{R}_{RAR})$ than in Γ. In addition, punishments after deviations, $\underline{W}_i(t)$, are also at least as small as the optimal punishment in game Γ because of the greater discounting which also affects the set of implementable outcomes at time t. That is, $\underline{W}_i(t) < \underline{W}_i$. Moreover, the fact that the implementability condition (5.5) depends on t also may violate the test devised in Proposition 3.1.

It turns out that in the long run (for t large enough), the implementability condition of sequences in Γ and $\widehat{\Gamma}(\boldsymbol{R}_{RAR})$ are arbitrarily close. This phenomenon is also mentioned in Charness and Genicot (2009). Specifically, we show in Lemma B.2 that

$$\lim_{t\to\infty} \frac{\eta(\beta, t+s+1)}{\eta(\beta, t)}\beta^s = \beta^s \text{ and } \lim_{t\to\infty}(1-\beta)\sum_{s=0}^{\infty}\frac{\eta(\beta,t+s+1)}{\eta(\beta,t)}\beta^s = 1.$$

As players keep on playing the IC constraints of the actual game and the induced game are not very different. Charness and Genicot (2009) take this as evidence that $\boldsymbol{R}_{RAR}$ "almost" implements Γ. We caution that this argument does not follow, for several reasons. First, as shown in the Prisoner's Dilemma example of section 2, repeated games may exhibit path dependence. Second, the speed of convergence of the incentive compatibility constraints is rather slow, as discussed below. Third, Charness and Genicot (2009) assess convergence in terms of the implicit discount rate. We argue that the correct measure of convergence ought to use the present value, which we show to converge at a much slower rate. Fourth, we prove that the participants will exhibit asymptotic indifference. Even though the IC constraints of the actual and induced games are asymptotically similar, agents simply will not care about what happens in any continuation game if t is high enough. The reason is simple: if a participant has been playing for a long enough time, whatever she does today only negligibly affects the expected value the lottery that she faces. Moreover, as she is discounting, the effect of future payoffs on the expected value of payments is also negligible. We formalize this idea below.

5.2.1. *Asymptotic Indifference.*

Definition 5.1 (Contribution)**.** Let $H : A^\infty \to \mathbb{R}$ be some function that has the property that it can be written as

$$H(\mathbf{a}) = \sum_{s=0}^{\infty} F(s, a_s)$$

for some function $F : \mathbb{N} \times A \to \mathbb{R}$. We define the *contribution of* a_s *to* H ,[7]

$$\mathcal{C}(H \mid s)(\mathbf{a}) := F(s, a_s).$$

Likewise, let $I \subset \mathbb{N}$ index set. We define the *contribution of* $\{a_k\}_{k\in I}$ as

$$\mathcal{C}(H \mid I)(\mathbf{a}) := \sum_{k\in I} F(k, a_k).$$

[7]The contribution of a_t to H is simply the Radón-Nikodym derivative of H with respect to the unit measure.

Example. Take $U_t^i(\mathbf{a}) = (1-\beta)\sum_{s=t}^{\infty}\beta^{s-t}u_i\left[r_i\left(a_s\right)\right]$ as the time t utility for agent i in game Γ at all subgames that start at date t. Then

$$\mathcal{C}\left(U_t^i \mid t\right)(\mathbf{a}) = (1-\beta)\, u_i\left[r_i\left(a_t\right)\right]$$

and

$$\mathcal{C}\left(U_t^i \mid s>t\right)(\mathbf{a}) = (1-\beta)\sum_{s=t+1}^{\infty}\beta^{s-t}u_i\left[r_i\left(a_s\right)\right]$$

and for the particular case of a stationary path (i.e. $u_i\left[r_i\left(a_s\right)\right] = \hat{u}_i$ for all s) then we can simplify the above as

$$\mathcal{C}\left(U_t^i \mid s>t\right)(\mathbf{a}) = \beta\hat{u}_i$$

Note that both are time independent and non-negligible.

Proposition 5.4 (Asymptotic Indifference)**.** *Let $\bar{u}^i = \max_{a\in A} u_i\left[r_i\left(a\right)\right]$. For all histories h^∞ and all $a \in A$,*

$$\mathcal{C}\left(U_t^i \mid t\right)(\mathbf{a}) \leq \bar{u}^i\,(1-\beta)\,\eta\left(\beta,t\right) = \bar{u}^i\frac{1}{t} + o\left(\frac{1}{t}\right) = o(1) \text{ as } t\to\infty,$$

and for all histories h^∞,

$$\mathcal{C}\left(U_t^i \mid s>t\right)(\mathbf{a}) \leq \bar{u}^i\,(1-\beta)\,\beta\eta_\beta\left(\beta,t\right) = \bar{u}^i\frac{\beta^2}{t\,(1-\beta)} + o\left(\frac{1}{t}\right) = o(1) \text{ as } t\to\infty.$$

5.2.2. *Measuring Distortions in Implementability and Payoffs.*
Relative Present Values in IC constraints. Equations (A.1) and (5.5) allow us to compare the implementability constraints quite easily. In each of the expressions, the present is evaluated in the same manner, $(1-\beta)\,u_i\left[r_i\left(a_t\right)\right]$, and the only differences come from discounting future payoffs, which is time dependent.

Suppose we consider a constant outcome $a_s = a$ for all $s \geq t$ at time t, which of course generates a constant stream of future utilities. The theoretical expected present value, which we will denote as W_t, is

$$W_t = (1-\beta)\sum_{s=0}^{\infty}\beta^s u = u.$$

On the other hand, the when we do this computation in the game $\widehat{\Gamma}\left(\boldsymbol{R}_{RAR}\right)$, we have

$$\widehat{W}_t = (1-\beta)\sum_{s=0}^{\infty}\left[\frac{\eta\left(\beta,t+s+1\right)}{\eta\left(\beta,t\right)}\right]\beta^s u = (1-\beta)\,\frac{\eta_\beta\left(\beta,t\right)}{\eta\left(\beta,t\right)}u$$

In turn, for any utility level u, we can define the *ratio of present values* ρ_t as

$$\rho_t := \frac{\widehat{W}_t}{W_t} = (1-\beta)\,\frac{\eta_\beta\left(\beta,t\right)}{\eta\left(\beta,t\right)}. \tag{5.6}$$

We know that $\rho_t \to 1$ and $\rho_t < 1$ for all t, since agents behave as if they discounted the future more than they actually do. With (5.6) we are equipped with an explicit measure of how bad the problem is. We can calculate the functions η and η_β very accurately, using the

finite integral formulation in 3 and 4 in Lemma B.1. This is a measure of the distortion in the implementability condition.

In Figure 5.2.2 we explore the behavior of this ratio for different values of β, as t grows.

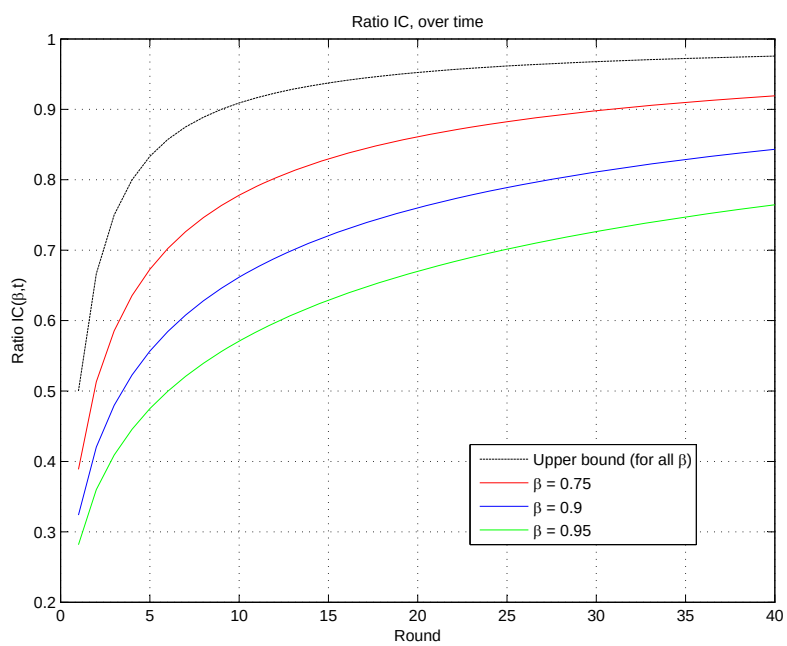

FIGURE 1. Presents the IC Ratio vs Time

We note that for $\beta = 0.9$, at round 1 the ratio is less than 0.3 and even by round 10 the ratio is still less than 0.6. In addition the figure displays a uniform bound across all discount factors. To demonstrate how the slow convergence relates to the asymptotic indifference, we introduce the relative contribution at time t as

$$\text{Relative contribution at } t := \frac{\mathcal{C}\left(U_t^i \mid s \geq t\right)}{\mathcal{C}\left(U_t^i \mid s < t\right) + \mathcal{C}\left(U_t^i \mid s \geq t\right)}.$$

This captures the share of one's utility which is comprised of today's and subsequent periods' decisions. In the target game notice that the relative contribution is identically 1 at all periods. By studying the behavior of the incentive compatibility ratio and the relative contribution together against time, we can see that as the IC slowly converges to the modeled IC, meanwhile the relative contribution rapidly converges to zero.

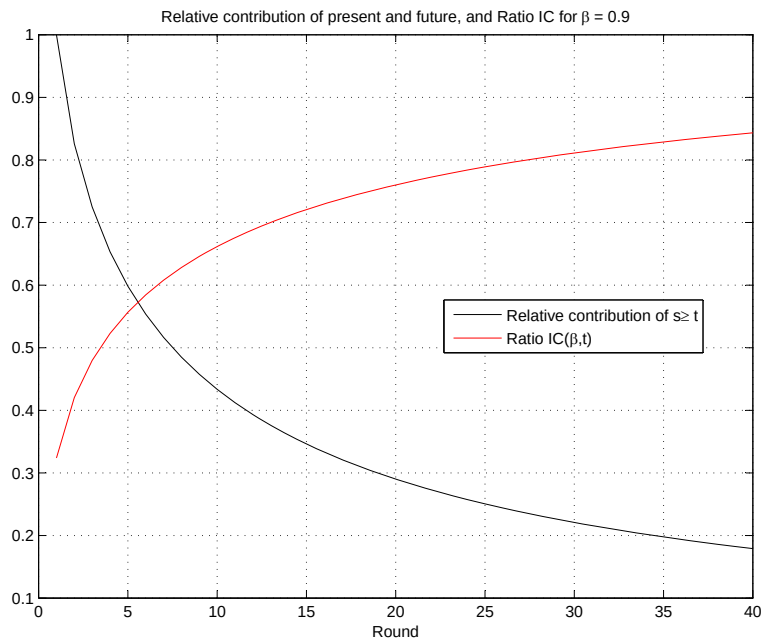

FIGURE 2. Presents the IC Ratio vs Time as well as the Relative Contribution vs Time

Note that by period 10, the relative contribution has dropped to 0.44 and the IC ratio is merely 0.67.

6. A Model of Savings

We observe that even though it is true that the set of implementable outcomes is distinct from the target even if the IC condition converges as $t \to \infty$, if agents can renegotiate at some point in the future, then in late rounds the outcomes in the induced game may resemble those of the target game. However, if we consider a dynamic game with a Markovian variable affected by past period interactions, this result would not hold. Namely, the potential renegotiation would be performed in a different subgame. To clarify this point we now consider the above framework in a common, restricted environment. An agent has an initial endowment of assets, $a_0 = 0$. At each t, the agent receives a deterministic endowment of $y_t \geq 0$ units and can save any amount at a constant interest rate $R > 0$. The budget constraint and no-Ponzi conditions are

$$c_t + a_{t+1} = y_t + Ra_t \text{ for all } t \in \mathbb{N} \text{ and } \lim_{t\to\infty} R^{-t} a_t = 0.$$

The agent has preferences given by $U = (1-\beta)\sum_{t=0}^{\infty} \beta^t u(c_t)$ where $u' > 0$ and $u'' < 0$. Arrow's time-zero constraint is

$$\sum_{t=0}^{\infty} R^{1-t} c_t = \sum_{t=0}^{\infty} R^{1-t} y_t \tag{6.1}$$

The usual Euler equation is $u'(c_t) = \beta R u'(c_{t+1})$. In particular $\beta R = 1$ yields $c_t = c_{t+1} = c^*$. The consequence of this, of course, is Friedman's permanent income hypothesis which follows by (6.1), with

$$\sum_{t=1}^{\infty} \beta^{t-1} c^* = \sum_{t=1}^{\infty} \beta^{t-1} y_t \Longleftrightarrow c^* = (1-\beta) \sum_{t=1}^{\infty} \beta^{t-1} y_t.$$

Note that this is independent of the specific preferences we considered.

6.1. **Round at Random Payment.** Suppose that we pay the agent for a randomly chosen round in the above environment. It can be shown, using the recursive method shown in Appendix B.2, that the modified Euler equation is

$$u'(c_t) = \underbrace{\eta(\beta,t)^{-1} \cdot \eta(\beta,t+1)}_{<1} \beta R_t u'(c_{t+1}) < \beta R u'(c_{t+1})$$

which immediately implies $c_t > c_{t+1}$. Instead of consuming a constant amount, the agent would choose a forever decreasing consumption bundle. To further our intuition, define $\hat{R}_t = \eta(\beta,t)^{-1} \cdot \eta(\beta,t+1) R_t$ as the effective gross interest rate, as the Euler equation is $u'(c_t) = \beta \hat{R}_t u(c_{t+1})$. Observe $\hat{R}_t < R$ for all periods; agents will save less under round at random payment than under the theoretical model. The following proposition shows the extent of this distortion.

Proposition 6.1. *Let $u(c)$ be a strictly concave and differentiable utility function, and allow interest rates to be time-varying (i.e. $R_t = 1 + r_t$ change over time). Then*

$$\frac{u'\left(c_t^{Theoretical}\right)}{u'\left(c_t^{RAR}\right)} = O\left(t^{-1}\right) \to 0 \text{ as } t \to \infty. \tag{6.2}$$

To illustrate the previous proposition, assume we have a CES utility function: $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$. Using the proposition we can show that

$$\frac{c_t^{\text{RAR}}}{c_t^{\text{Theoretical}}} \propto (\eta(\beta,t))^{\frac{1}{\sigma}} = O\left(t^{-\frac{1}{\sigma}}\right).$$

This implies that consumption under the RAR payment is infinitely smaller than the theoretical predicted consumption decision when $t \to \infty$. This seems to be at odds with the result that the implementability condition of RAR converges to the theoretical IC; given that in this "game" there is only one player, the agent should "renegotiate" to get the most efficient decision at any point in time. To see this more explicitely, suppose that $\beta R = 1$, so that theoretical consumption is constant, and moreover suppose $\sigma = 1$ (so $u(c) = \log(c)$). Then,

$$\frac{c_{t+1}}{c_t} = \frac{\eta(\beta,t+1)}{\eta(\beta,t)} \text{ and } c_t = \frac{\eta(\beta,t)}{\eta(\beta,1)} c_1.$$

Since $\eta(\beta,t) \to 0$, clearly $c_t \to 0$. Even if the IC constraint does converge to the one of the theoretical model (since $\eta(\beta,t)^{-1} \cdot \eta(\beta,t+1) \to 1$ as $t \to \infty$) the behavior of the solution

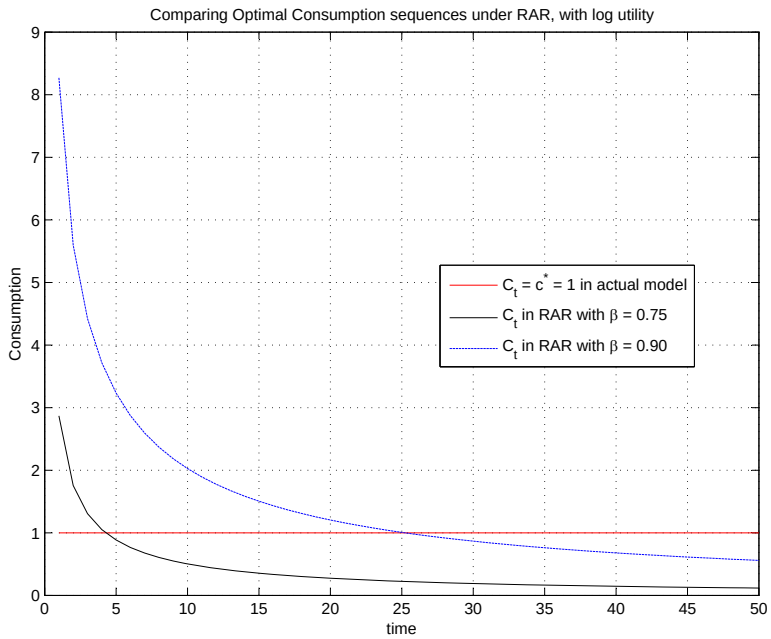

FIGURE 3. Consumption Sequences with log utility for RAR.

does not approximate the one in the theoretical model. In particular, the solution at very large t does resemble a constant consumption, but the wrong constant. Instead of $c_t = c^*$, it will become arbitrarily close to zero. Normalizing $c^* = 1$ we compare the consumption sequences under the two models in Figure 6.1, under different values of β.

In early rounds agents consume more than prescribed by the theoretical model. In particular, for $\beta = 0.9$ consumption at $t = 1$ is 8.26 times bigger than consumption in the first period and at $t = 2$ is 5.6 times bigger, for any gross interest rate $R > 1$.

6.2. **Payments for All Rounds.** We now consider the case where agents are paid for all rounds. In Appendix B.3 we show that the modified Euler equation is now

$$u'(c_t) = \beta (R_t - 1) u'(c_{t+1}) < \beta R_t u'(c_{t+1}). \tag{6.3}$$

Again this has the effect of reducing incentives for saving, since the effective gross interest rate for the agent is $\hat{R}_t := R_t - 1$. The extent of the distortion is illustrated by the following proposition.

Proposition 6.2. *Let $u(c)$ be a strictly concave and differentiable utility function, and allow interest rates to be time-varying (i.e. $R_t = 1 + r_t$ change over time). Then*

$$\frac{u'\left(c_t^{Theoretical}\right)}{u'\left(c_t^{All\ Rounds}\right)} \propto \prod_{s=1}^{t-1} \frac{r_s}{1+r_s} \to 0 \text{ as } t \to \infty. \tag{6.4}$$

In particular, if $r_s = r$ for all s, then

$$\frac{u'\left(c_t^{Theoretical}\right)}{u'\left(c_t^{All\ Rounds}\right)} = O\left(\left(1+\frac{1}{r}\right)^{-t}\right). \tag{6.5}$$

To illustrate this proposition, we return to the CES example. Notice

$$\frac{c_t^{\text{All Rounds}}}{c_t^{\text{Theoretical}}} \propto \left(\frac{R-1}{R}\right)^{\frac{t-1}{\sigma}} \to 0 \text{ as } t \to \infty,$$

so no matter how small σ is the ratio between consumptions goes to zero. In particular, consider the case where $1 < R < 1 + \frac{1}{\beta}$. This implies that $\beta R > 1$ and $\beta(R-1) < 1$. The above equations imply that for any $\sigma > 0$ we have that

$$c_t^{\text{All Rounds}} \propto (\beta(R-1))^{\frac{t-1}{\sigma}} \to 0 \text{ since } \beta(R-1) < 1$$

but

$$c_t^{\text{Theoretical}} \propto (\beta R)^{\frac{t-1}{\sigma}} \to \infty \text{ since } \beta R > 1.$$

Therefore, we have shown not only that the ratio of consumption goes to zero but also that the behavior of the optimal solution path is extremely different. Furthermore, notice that this is true no matter how much curvature (i.e. $\sigma > 0$) we assume. We highlight this point because we know that when utility is linear (i.e. $\sigma = 0$) payment in all rounds does implement the actual game. However, this result is largely non-generic: allowing for arbitrarily small amounts of curvature (in the CES family) implementable outcomes are starkly different.

To illustrate numerically this feature, we set $\beta = 0.97$, $\sigma = \frac{1}{2}$ and $R = 1.2$.[8] Normalizing the net present value of endowments so that $c_1^* = 1$, we present the results in Figure 6.2.

For $\beta R = 1$ we have that $c_t > c_{t+1}$, so consumption is decreasing even when c_t should be constant, similar to what we observed under the RAR payment scheme. Using log-utility, it is easy to see that

$$\frac{c_{t+1}}{c_t} = \beta(R-1) = 1 - \beta \text{ if } R\beta = 1.$$

For β used in the litereature, typically close to one, this is a rapidly decaying function. Figure 6.2 presents a numerical illustration.

Intuitively, an agent who is very patient (thereby facing a low interest rate) can essentially consume as much as possible in the first period, by borrowing against the flow of future income. Since all future streams of consumption until the end of the game is kept by the

[8]Smaller values of R or σ make results much starker.

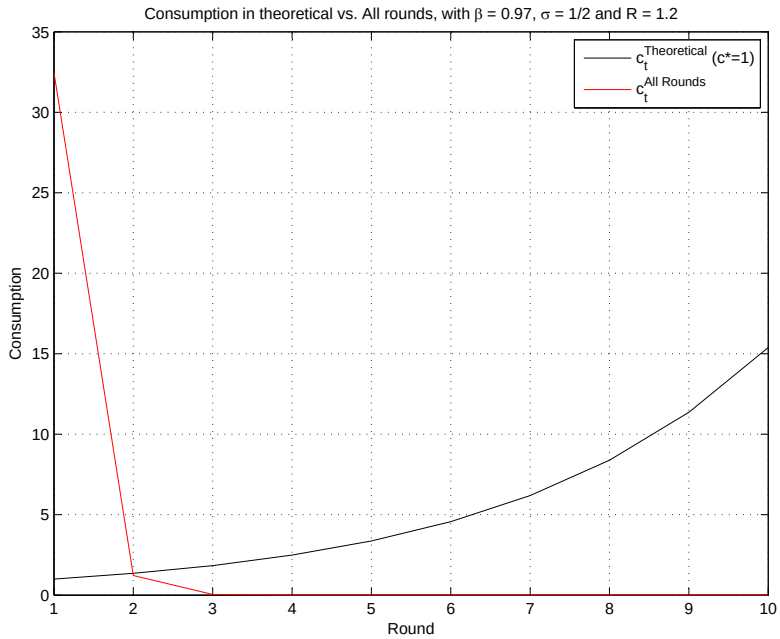

FIGURE 4. Consumption in theoretical model versus all round payment

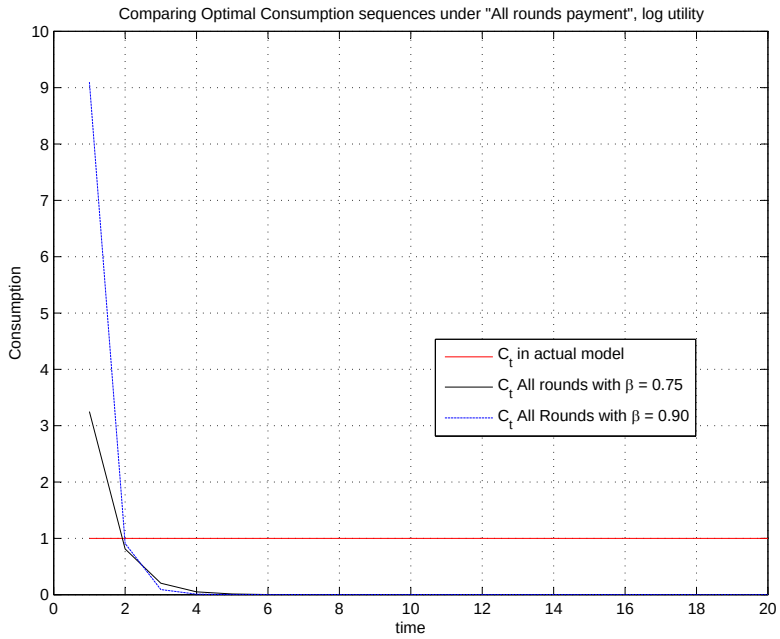

FIGURE 5. Consumption Sequences with log utility for All Round Payment.

agent, the she wants to maximize the amount of consumption early on. Let $c^* = 1$ and observe that

$$c_1 = \sum_{t=1}^{\infty} \beta^{t-1} y_t = \frac{1}{1-\beta} c^* = \frac{1}{1-\beta} \to \infty \text{ as } \beta \to 1$$

Turning to Figure 6.2, when $\beta = 0.9$ the agent consumes $c_1 = 9.1$ while the upper bound of consumption is $\frac{1}{1-0.9} = 10$. This has to be followed by a steep reduction in consumption, as seen in the figure, in order to be able to repay it. When $\beta = 0.99$ the participant gets even closer: consumption in the first period is $c_1 = 99.01$ while $\frac{1}{1-0.99} = 100$.

7. Conclusion

We have discussed payment schemes in repeated games. When the researcher is interested in implementing a particular game, the payment schemes often used in the literature fail. In particular, we have characterized how the implementability conditions are changed by the payment schemes. We add that since our characterizations are analytically explicit, one can in principle use the modified implementability conditions to then analyze the behavior in data generated from these payment schemes. In addition, we have found a simple payment scheme which implements a much broader class of games by simply paying participants for the last round of play.

References

Abreu, D. (1988): "On the theory of infinitely repeated games with discounting," *Econometrica: Journal of the Econometric Society*, 383–396.

Aoyagi, M. and G. Fréchette (2009): "Collusion as public monitoring becomes noisy: Experimental evidence," *Journal of Economic Theory*, 144, 1135–1165.

Bo, P. (2005): "Cooperation under the shadow of the future: experimental evidence from infinitely repeated games," *The American economic review*, 95, 1591–1604.

Charness, G. and G. Genicot (2009): "Informal Risk Sharing in an Infinite-Horizon Experiment*," *The Economic Journal*, 119, 796–825.

Duffy, J. and J. Ochs (2009): "Cooperative behavior and the frequency of social interaction," *Games and Economic Behavior*, 66, 785–812.

Engle-Warnick, J. and R. Slonim (2004): "The evolution of strategies in a repeated trust game," *Journal of Economic Behavior & Organization*, 55, 553–573.

——— (2006a): "Inferring repeated-game strategies from actions: evidence from trust game experiments," *Economic Theory*, 28, 603–632.

——— (2006b): "Learning to trust in indefinitely repeated games," *Games and Economic Behavior*, 54, 95–114.

Fischer, G. (2010): "Contract structure, risk sharing, and investment choice," *February. Available at: http://personal. lse. ac. uk/fischerg/Research. htm.*

Kreps, D., P. Milgrom, and J. RobertsRobert (1982): "Rational cooperation in the finitely repeated prisoners' dilemma," *Journal of Economic Theory*, 27, 245–252.

Palfrey, T. and H. Rosenthal (1994): "Repeated play, cooperation and coordination: An experimental study," *The Review of Economic Studies*, 61, 545.

Appendix A. Proofs

Proof of Proposition 3.1. We need to show that, conditional on the future play of the game, the set of implementable strategy profiles at time t is independent of the past play of the game; i.e $O_t\left(\widehat{\Gamma}\left(\boldsymbol{R}\right), h^{t-1}|\bar{\mathbf{a}}'\right)$. First, we will show a simple, yet important property. Take two equilibrium profiles $\sigma_1$and σ_2 with the property that these strategies predict the same equilibrium path from some history onwards. That is, there exist histories $h_1^{t_1-1}$and $h_2^{t_2-1}$ such that $a_{t_1+s}\left(\sigma_1|h_1^{t_1-1}, a_{t_1}\left(\sigma_1|h_1^{t_1-1}\right)\right) = a_{t_2+s}\left(\sigma_2|h_2^{t_2-1}, a_{t_2}\left(\sigma_1|h_2^{t_2-1}\right)\right) = \bar{a}_s$ for all $s > 1$. We will show that in this case, the set of admissible equilibrium action profiles at time t_1 for $\sigma_1$and at time t_2 for σ_2 should coincide. That is, $O_{t_1}\left(\Gamma, h_1^{t_1-1} \mid \{a_{t_1+s}\}_{s=1}^{\infty}\right) = O_{t_2}\left(\Gamma, h_2^{t_2-1} \mid \{a_{t_2+s}\}_{s=1}^{\infty}\right)$. This implies that the past history should be irrelevant to predict the current equilibrium profile, conditional on future play under the equilibrium profile.

To simplify notation, we take only one of the equilibrium profiles and drop the indices $j \in \{1, 2\}$. Because $A = \prod_{i=1}^{i=n} A^i$ is compact and $u_i\left(r_i\left(\cdot\right)\right)$ are continuous functions of $a \in A$, we can use Proposition 5 in Abreu (1988) to guarantee the existence of a simple SPE profile $\hat{\sigma}$ that implements the same equilibrium path $a\left(\sigma\right)$. To characterize implementable outcomes we can restrict ourselves to simple profiles $\hat{\sigma}$ that includes optimal punishments for deviators. Take history h^{t-1} such that agents have been playing according to $\hat{\sigma}$ so far and consider a potential deviation from the equilibrium path by agent i, which we will denote as $\hat{a}_t^i := a_t^i\left(\hat{\sigma}|h^{t-1}\right)$. The one-shot deviation principle gives us an IC constraint that characterizes the equilibrium profile $\hat{\sigma}$:

$$\max_{\tilde{a}^i \in A^i} \left(1-\beta\right) u_i\left(r_i\left(\tilde{a}^i, \hat{a}_t\right)\right) + \beta \underline{W}^i \leq \left(1-\beta\right) u_i\left(r_i\left(\hat{a}_t\right)\right) + \beta \sum_{t=0}^{\infty} \left(1-\beta\right) \beta^s u_i\left(r_i\left(\bar{a}_s\right)\right),$$

where $\underline{W}^i$ is player i's optimal punishment payoff, which is history independent. Equivalently, we may write this as

$$\hat{a}_t^i \in \operatorname*{argmax}_{\tilde{a}^i \in A^i} \left(1-\beta\right) u_i\left(r_i\left(\tilde{a}^i, \hat{a}_t\right)\right) + \beta W^i\left(\tilde{a}^i, \hat{a}_t\right) \tag{A.1}$$

where

$$W^i\left(\tilde{a}^i, \hat{a}_t\right) = \begin{cases} \underline{W}^i & \text{if} \quad \tilde{a}^i \neq \hat{a}_t^i \\ \left(1-\beta\right) \sum_{s=0}^{\infty} \beta^s \beta^s u_i\left(r_i\left(\bar{a}_s\right)\right) & \quad \text{otherwise} \end{cases}.$$

Conditional on having reached history h^{t-1}, the maximized function in (A.1) is independent of the past play. Assuming we are playing according to a SPE profile, the IC constraint in (A.1) must be independent of the past play, which implies that if two different simple SPE profiles have the same continuation equilibria, then the past history is completely irrelevant to set the maximizer of (A.1), proving the desired claim (using σ_1 and σ_2 to establish the result). The fact that $\boldsymbol{R}$ implements Γ proves then that given a future path of play □

Proof of Proposition 4.1. The proof is straightforward. Fix some history h^{t-1} and some SPE $\sigma\left(h^{t-1}\right)$. Continuation utilities for any profile (a_t) given scheme $\boldsymbol{R}_{last}$ are

$$W_{i,t}\left(a_t\right)=\left(1-\beta\right)\sum_{s=1}^{\infty}\beta^s u_i\left[r_i\left(a_{t+s}\left(\sigma\right)\right)\right]$$

which coincides with the expected payment under $\boldsymbol{R}_{last}$ if agents play according to σ, for any profile a_t. So, for any SPE profile σ, the one-shot deviation principle states that

$$a_{i,t}\in\underset{\hat{a}^i\in A^i}{\operatorname{argmax}}\left(1-\beta\right)u^i\left(r_i\left(\hat{a}^i,a_t^{-i}\right)\right)+\beta W_{i,t}\left(\hat{a}^i,a_t^{-i}\right)$$

which coincides with the one-shot deviation principle of game $\widehat{\Gamma}\left(\boldsymbol{R}_{last}\right)$, since the game ends with probability$(1-\beta)$and continues with probability β. Therefore, profile σ is also an SPE profile of $\widehat{\Gamma}\left(\boldsymbol{R}_{last}\right)$. The reverse is trivial. □

Proof of Proposition 5.2. Observe that

$$U_i\left(\mathbf{a}\mid h^t\right)-U_i\left(\mathbf{a}'\mid h^t\right)=\left(1-\beta\right)\sum_{s=0}^{\infty}\beta^s\left[u_i\left(R_{i,t}+\sum_{k=1}^{s}r_{t+k}\right)-u_i\left(R_{i,t}+\sum_{k=1}^{s}r'_{t+k}\right)\right]$$

where $r_{i,t+k}=r_i\left(a_{t+k}\right)$ and likewise for $r'_{i,t+k}$. By concavity and differentiability of $u_i(.)$, we know that for all $x,y\in\mathbb{R}$ we have that

$$u'\left(x\right)\left(x-y\right)\leq u\left(x\right)-u\left(y\right)\leq u'\left(y\right)\left(x-y\right).$$

Using $x=R_{i,t}+\sum_{k=1}^{s}r_{t+k}$ and $y=R_{i,t}+\sum_{k=1}^{s}r'_{t+k}$,

$$\begin{aligned}\left(1-\beta\right)\sum_{s=0}^{\infty}\beta^s u'_i\left(R_{i,t}+\sum_{k=1}^{s}r_{i,t+k}\right)\left(\sum_{k=1}^{s}\left(r_{i,t+k}-r'_{i,t+k}\right)\right) &\leq U_i\left(\mathbf{a}\mid h^t\right)-U_i\left(\mathbf{a}'\mid h^t\right)\\ &\leq \left(1-\beta\right)\sum_{s=0}^{\infty}\beta^s u'_i\left(R_{i,t}+\sum_{k=1}^{s}r'_{i,t+k}\right)\left(\sum_{k=1}^{s}\left(r_{i,t+k}-r'_{i,t+k}\right)\right)\end{aligned}$$

As u_i is concave, $u'_i\left(R_{i,t}+\sum_{k=1}^{s}r_{i,t+k}\right)\leq u'\left(R\right)$ and the same for $r'_{i,t+k}$, so

$$\left|U_i\left(\mathbf{a}\mid h^t\right)-U_t^i\left(\mathbf{a}'\mid h^t\right)\right|\leq u'\left(R_{i,t}\right)\left(1-\beta\right)\sum_{s=0}^{\infty}\beta^s\left|r_{i,t+k}-r'_{i,t+k}\right|$$

If $r_i^{\min}>0$, then $R_{i,t}\geq\left(t+1\right)r_i^{\min}$, implying that

$$\left|U_i\left(\mathbf{a}\mid h^t\right)-U_t^i\left(\mathbf{a}'\mid h^t\right)\right|<u'_i\left[\left(t+1\right)r_i^{\min}\right]\left[\left(1-\beta\right)\sum_{s=0}^{\infty}\beta^s\left|r_i^{\max}-r_i^{\min}\right|\right]\leq u'_i\left[\left(t+1\right)r_i^{\min}\right]\left(r_i^{\max}-r_i^{\min}\right).$$

as we wanted to show. To show the contemporary utility result, without loss of generality take $r,\widehat{r}\in r_i\left(A\right)$, such that $r>\widehat{r}$. Then,

$$\begin{aligned}u'_i\left(R_{i,t}+r_i^{\max}\right)\left(r-\widehat{r}\right) &\leq u'_i\left(R_{i,t}+r\right)\left(r-\widehat{r}\right)\leq\left(1-\beta\right)u_i\left(R_{i,t}+r\right)-\left(1-\beta\right)u_i\left(R_{i,t}+\widehat{r}\right)\\ &\leq u'_i\left(R_{i,t}+\widehat{r}\right)\left(r-\widehat{r}\right)\leq u'_i\left(R_{i,t}+r_i^{\max}\right)\left(r_i^{\max}-r_i^{\min}\right)\end{aligned}$$

which implies that

$$|u_i(R_{i,t}+r)-u_i(R_{i,t}+\hat{r})| \leq u_i'\left(R_{i,t}+r_i^{\min}\right)\left(r_i^{\max}-r_i^{\min}\right) \leq u_i'\left[(t+1)\,r_i^{\min}\right]\left(r_i^{\max}-r_i^{\min}\right).$$

as we wanted to show. □

Proof of Proposition 5.3. (Sketch) The proposition is a direct consequence of Lemma B.4: $\{a_t\}$ is an implementable outcome iff is the outcome of some SPE strategy σ, and such strategy has to be such that $\{a_{t+s}\}_{s=0}^{\infty} \in \mathbf{O}(\Gamma_t)$ (i.e. is a SPE outcome of the game starting at sub-tree h^t, which is identical to Γ_t, up to difference in intercepts in the utility functions), which we showed in Lemma B.4, together with the fact that it is continuous at infinity, and has a compact $V(\Gamma_t)$. Then, following the same reasoning as in Abreu (1988) we conclude that we can implement any outcome by using optimal simple punishments, which has payoff $\underline{W^i}(t)$. Continuation values in Γ_t are calculated using the discount function $B(\beta, t+1, s)$ for all $s \geq 1$, finishing the proof. □

Proof of Proposition 5.4. From Lemma B.3, equation B.4, we have that for any a_t the contribution of t period strategies is $\mathcal{C}\left(U_t^i \mid t\right)(a_t) = (1-\beta)\,\eta(\beta,t)\,u_i\left[r_i(a_t)\right]$ and the result follows from there. Also, Property 5 in Lemma B.1 gives the approximation result (and convergence to zero). For any history a_{t+s}, using again equation B.4, we have

$$\begin{aligned}
\mathcal{C}\left(U_t^i \mid \{s \geq t+1\}\right)\left(\{a_{t+s}\}_{s=1}^{\infty}\right) &= \beta(1-\beta)\sum_{s=0}^{\infty}\eta(\beta,t+1+s)\,u_i\left[r_i(a_{t+1+s})\right] \\
&\leq \beta\bar{u}_i(1-\beta)\sum_{s=0}^{\infty}\sum_{k=s}^{\infty}\frac{\beta^k}{t+1+k} = \beta\bar{u}_i(1-\beta)\sum_{k=0}^{\infty}\sum_{s=0}^{k}\frac{\beta^k}{t+1+k} \\
&= \beta\bar{u}_i(1-\beta)\sum_{k=0}^{\infty}\frac{(k+1)\,\beta^k}{t+(k+1)} = \beta\bar{u}_i(1-\beta)\left(\sum_{j=1}^{\infty}\frac{j\beta^{j-1}}{t+j}\right) \\
&= \beta\bar{u}_i(1-\beta)\left(\sum_{j=0}^{\infty}\frac{j\beta^{j-1}}{t+j}-0\right) = \beta\bar{u}_i(1-\beta)\,\eta_\beta(\beta,t)
\end{aligned}$$

and the approximation result comes from property 6 in Lemma B.1. □

Proof of 6.1. Using the Euler equation for both the theoretical and the RAR sequence of optimal consumption, we get

$$u'\left(c_t^{\text{Theoretical}}\right) = \frac{1}{\beta^{t-1}\left(\prod_{s=1}^{t-1}R_s\right)}u'\left(c_1^{\text{Theoretical}}\right)$$

and

$$u'\left(c_t^{\text{RAR}}\right) = \frac{1}{\beta^{t-1}\left(\prod_{s=1}^{t-1} R_s\right)} u'\left(c_1^{\text{RAR}}\right) \frac{\eta(\beta,1)}{\eta(\beta,t)}$$

then

$$\frac{u'\left(c_t^{\text{Theoretical}}\right)}{u'\left(c_t^{\text{RAR}}\right)} = \frac{u'\left(c_1^{\text{Theoretical}}\right)}{u'\left(c_1^{\text{RAR}}\right)} \frac{\eta(\beta,t)}{\eta(\beta,1)} \propto \eta(\beta,t) = O\left(t^{-1}\right)$$

since $\eta(\beta,t) = \frac{1}{t(1-\beta)} + o\left(t^{-1}\right)$. □

Proof. From before we knew that

$$u'\left(c_t^{\text{Theoretical}}\right) = \frac{1}{\beta^{t-1}\prod_{s=1}^{t-1} R_s} u'\left(c_1^{\text{Theoretical}}\right)$$

and likewise, we show that

$$u'\left(c_t^{\text{All Rounds}}\right) = \frac{1}{\beta^{t-1}\prod_{s=1}^{t-1} (R_s - 1)} u'\left(c_1^{\text{All Rounds}}\right)$$

Therefore

$$\frac{u'\left(c_t^{\text{Theoretical}}\right)}{u'\left(c_t^{\text{All Rounds}}\right)} = \frac{\frac{1}{\beta^{t-1}\prod_{s=1}^{t-1} R_s} u'\left(c_1^{\text{Theoretical}}\right)}{\frac{1}{\beta^{t-1}\prod_{s=1}^{t-1} (R_s-1)} u'\left(c_1^{\text{All Rounds}}\right)} \propto \prod_{s=1}^{t-1} \frac{r_s}{1+r_s}$$

if $r_s = r$ then is clear that $\prod_{s=1}^{t-1} \frac{r_s}{1+r_s} = \left(1+\frac{1}{r}\right)^{-t+1} = O\left(\left(1+\frac{1}{r}\right)^{-t}\right)$ □

Appendix B. Auxiliary Results

B.1. Technical Results.

Lemma B.1. *Let η and ϕbe defined as in 5.2 and let $\eta_\beta(\beta,t) := \frac{\partial \eta(\beta,t)}{\partial \beta}$. Then*

(1) For all $\beta \in (0,1)$ and all $t \geq 1$ we have $\eta(\beta,t+1) = \frac{1}{\beta}\left(\eta(\beta,t) - \frac{1}{t}\right)$
(2) For $t \geq 1$ we can write $\eta(\beta,t)$ as $\eta(\beta,t) = \frac{1}{\beta^t}\ln\left(\frac{1}{1-\beta}\right) - \sum_{i=1}^{t-1} \frac{1}{(t-i)\beta^i}$
(3) For all β, t we can write $\eta(\beta,t) = \frac{1}{\beta}\int_0^\beta \left(\frac{z}{\beta}\right)^t \frac{1}{z(1-z)} dz$
(4) For all β, t we can write η_β as $\eta_\beta(\beta,t) = \sum_{k=0}^\infty \frac{k\beta^{k-1}}{t+k} = \int_0^\beta \left(\frac{z}{\beta}\right)^t \frac{1}{(1-z)^2} dz$
(5) For givenβ: $\eta(\beta,t) = \frac{1}{t(1-\beta)} + o\left(\frac{1}{t}\right)$ and therefore $\lim_{t\to\infty} \eta(\beta,t) = 0$
(6) For all β, $\eta_\beta(\beta,t) = \frac{\beta}{t(1-\beta)^2} + o\left(\frac{1}{t}\right)$, and therefore $\lim_{t\to\infty} \eta_\beta(\beta,t) = 0$
(7) For all β, t, s we have that $\sum_{k=s}^\infty \frac{\beta^k}{t+k} = \beta^s \eta(\beta,t+s)$

Proof. We first prove 2. This proof is by induction. For $t = 1$ we have that

$$\eta(\beta, 1) = \sum_{k=0}^{\infty} \frac{\beta^k}{1+k}$$

it can be shown, using integration and Abel's Theorem, that

$$\eta(\beta, 1) = \frac{1}{\beta} \log\left(\frac{1}{1-\beta}\right)$$

To prove it, we need to prove the following recursion

$$\eta(\beta, t+1) = \frac{1}{\beta}\left[\eta(\beta, t) - \frac{1}{t}\right] \tag{B.1}$$

and is easy to see that η as defined in 5.2 satisfies this recursion. To show that this recursion is true, we do some algebra:

$$\eta(\beta, t+1) = \sum_{s=0}^{\infty} \frac{\beta^s}{t+(1+s)} = \sum_{j=1}^{\infty} \frac{\beta^{j-1}}{t+j} = \frac{1}{\beta}\left[\sum_{j=0}^{\infty} \frac{\beta^j}{t+j} - \frac{\beta^0}{t+0}\right] = \frac{1}{\beta}\left[\eta(\beta, t) - \frac{1}{t}\right]$$

For 3, see that

$$\sum_{k=0}^{\infty} \beta^k = \frac{1}{1-\beta} \Longrightarrow \sum_{k=0}^{\infty} \beta^{t+k-1} = \frac{\beta^{t-1}}{1-\beta} \Longrightarrow \sum_{k=0}^{\infty} \frac{\beta^{t+k}}{t+k} = \int_0^{\beta} \frac{z^{t-1}}{(1-z)} dz \Longleftrightarrow$$

$$\eta(\beta, t) = \frac{1}{\beta} \int_0^{\beta} \left(\frac{z}{\beta}\right)^t \frac{1}{1-z} dz$$

which is valid since η is a power series. For 4 we can also use this to differentiate η:

$$\eta_\beta(\beta, t) = \frac{\partial}{\partial \beta}\left(\sum_{k=0}^{\infty} \frac{\beta^k}{t+k}\right) = \sum_{k=0}^{\infty} \frac{k\beta^{k-1}}{t+k}$$

also

$$\sum_{k=0}^{\infty} k\beta^{k-1} = \frac{d\left(\frac{1}{1-\beta}\right)}{d\beta} = \frac{1}{(1-\beta)^2} \Longrightarrow \sum_{k=0}^{\infty} k\beta^{t+k-1} = \frac{\beta^t}{(1-\beta)^2} \Longleftrightarrow$$

$$\eta_\beta(\beta, t) = \sum_{k=0}^{\infty} \frac{k\beta^{k-1}}{t+k} = \int_0^{\beta} \left(\frac{z}{\beta}\right)^t \frac{1}{(1-z)^2} dz$$

For 5, we must show that $t\eta(\beta, t) \to \frac{1}{1-\beta}$ as $t \to \infty$. We can write $t\eta(\beta, t) = \sum_{k=0}^{\infty} \frac{t}{t+k}\beta^k$. Defining the sequence of sequences $f_t(k) := \frac{t}{t+k}\beta^k$ is easy to see that $f_t \nearrow \beta^k$ pointwise. Therefore, we can use the Dominated convergence theorem to show that $\lim_{t\to\infty} t\eta(\beta, t) = \sum_{k=0}^{\infty}\left(\lim_{t\to\infty} \frac{t}{t+k}\beta^k\right) = \frac{1}{1-\beta}$. The convergence to 0 of η is straightforward and omitted.

For 6 we follow the same strategy, and note that $t\eta_\beta(\beta, t) = \sum_{k=0}^{\infty} \frac{tk}{t+k}\beta^k$. We have that $\frac{tk}{t+k}\beta^k \nearrow k\beta^k$ pointwise, which implies that $\lim_{t\to\infty} t\eta_\beta(\beta, t) = \sum_{k=0}^{\infty} k\beta^k = \frac{\beta}{(1-\beta)^2}$

Finally, for 7 see that $\phi(\beta, t, s) = \sum_{k=s}^{\infty} \frac{\beta^k}{t+k} = \sum_{j=0}^{\infty} \frac{\beta^{s+j}}{t+s+j} = \beta^s \eta(\beta, t+s)$ □

Lemma B.2. *Define the function $B(\beta,t,s)$, as*

$$B(\beta,t,s)=\beta^s\frac{\eta(\beta,t+1+s)}{\eta(\beta,t)}. \tag{B.2}$$

Then, the following hold:

(1) $B(\beta,t,s)<\beta^s$ *for all* $t,s\in\mathbb{N}$.
(2) $B(\beta,t,s)$ *is increasing in* t *and decreasing in* s.
(3) $(1-\beta)\sum_{s=0}^{\infty}B(\beta,t,s)=(1-\beta)\frac{\eta_\beta(\beta,t)}{\eta(\beta,t)}\to 1$ *as* $t\to\infty$.
(4) $\lim_{s\to\infty}B(\beta,t,s)=0$ *for all* $t\in\mathbb{N}$.
(5) $\lim_{t\to\infty}B(\beta,t,s)=\beta^s$, *so*

$$B(\beta,t,s)\nearrow\beta^s \text{ for all } s, \text{ as } t\to\infty. \tag{B.3}$$

Proof. (1) is obvious, since $\eta(\beta,t)$is decreasing in t.We first show (5). See that we can write

$$B(\beta,t,s)=\frac{\sum_{k=s}^{\infty}\frac{\beta^k}{t+1+k}}{\sum_{k=0}^{\infty}\frac{\beta^k}{t+k}}=\frac{\sum_{k=s}^{\infty}\left(\frac{t}{t+1+k}\right)\beta^k}{\sum_{k=0}^{\infty}\left(\frac{t}{t+k}\right)\beta^k}$$

so

$$\lim_{t\to\infty}B(\beta,t,s)\underset{(i)}{=}\frac{\sum_{k=s}^{\infty}\left(\lim_{t\to\infty}\frac{t}{t+1+k}\beta^k\right)}{\sum_{k=0}^{\infty}\left(\lim_{t\to\infty}\frac{t}{t+k}\beta^k\right)}=\frac{\sum_{k=s}^{\infty}\beta^k}{\sum_{k=0}^{\infty}\beta^k}=$$
$$=\frac{\left(\frac{\beta^s}{1-\beta}\right)}{\left(\frac{1}{1-\beta}\right)}=\beta^s$$

Using in (i) the Uniform Convergence theorem (the summand sequences are monotone decreasing in k). Moreover, is easy to show (with some tedious algebra) that B is decreasing in t and increasing in s (proving (2)). Facts (2) with (5) implies (1). That $B(\beta,t,s)\to 0$ as $s\to\infty$ follows directly from the fact that $\sum_{k=s}^{\infty}\frac{1}{t+k}\beta^k\to 0$ as $s\to\infty$. Finally,

$$\sum_{s=0}^{\infty}B(\beta,t,s)=\sum_{s=0}^{\infty}\frac{\sum_{k=s}^{\infty}\left(\frac{1}{t+1+k}\right)\beta^k}{\eta(\beta,t)}=\frac{1}{\eta(\beta,t)}\sum_{s=0}^{\infty}\sum_{k=s}^{\infty}\left(\frac{1}{t+1+k}\right)\beta^k=$$
$$\frac{1}{\eta(\beta,t)}\sum_{k=0}^{\infty}\sum_{s=0}^{k}\left(\frac{1}{t+1+k}\right)\beta^k=\frac{1}{\eta(\beta,t)}\sum_{k=0}^{\infty}\frac{(k+1)\beta^k}{t+1+k}=\frac{\eta_\beta(\beta,t)}{\eta(\beta,t)}$$

□

Lemma B.3. *After history h^t, time t, the utility of agent i of outcome $\mathbf{a}=\{a_t\}_{t=0}^{\infty}$ in $\hat{\Gamma}(\boldsymbol{R}_{RAR})$ can be written as*

(B.4)
$$\hat{U}_i\left(\mathbf{a}\mid h^t\right)=(1-\beta)\eta(\beta,t)\sum_{k=0}^{t}u_i\left[r_i(a_k)\right]+\beta\left\{(1-\beta)\sum_{s=0}^{\infty}\beta^s\eta(\beta,t+1+s)u_i\left[r_i(a_{t+s+1})\right]\right\}$$

Proof. We can decompose the utility as

$$U_t^i = (1-\beta)\frac{1}{t}\sum_{k=0}^{t} u_i\left[r_i\left(a_t\right)\right] + \beta W^i \tag{B.5}$$

, where $W^i$is the discounted present value of future periods payoffs. We can calculate it as

$$W^i = (1-\beta)\sum_{j=0}^{\infty}\beta^j\frac{1}{t+1+j}\left(\sum_{k=0}^{t} u_i\left[r_i\left(a_k\right)\right] + \sum_{k=0}^{j} u^i\left[r_i\left(a_{t+k+1}\right)\right]\right) =$$

$$(1-\beta)\sum_{j=0}^{\infty}\beta^j\frac{1}{t+1+j}\sum_{k=0}^{t} u_i\left[r_i\left(a_t\right)\right] + (1-\beta)\sum_{j=0}^{\infty}\beta^j\frac{1}{t+1+j}\sum_{k=0}^{j} u_i\left[r_i\left(a_{t+k+1}\right)\right] =$$

$$(1-\beta)\,\eta\left(\beta,t+1\right)\sum_{k=0}^{t} u_i\left[r_i\left(a_t\right)\right] + (1-\beta)\sum_{j=1}^{\infty}\sum_{k=0}^{j}\frac{\beta^j}{t+1+j}u_i\left[r_i\left(a_{t+k+1}\right)\right] =$$

$$(1-\beta)\,\eta\left(\beta,t+1\right)\left(\sum_{k=0}^{t} u_i\left[r_i\left(a_t\right)\right]\right) + (1-\beta)\sum_{k=0}^{\infty} u_i\left[r_i\left(a_{t+k+1}\right)\right]\left(\sum_{j=k}^{\infty}\frac{\beta^j}{t+1+j}\right) \Longleftrightarrow$$

$$W^i = (1-\beta)\,\eta\left(\beta,t+1\right)\left(\sum_{k=0}^{t} u_i\left[r_i\left(a_t\right)\right]\right) + (1-\beta)\sum_{s=0}^{\infty}\beta^s\eta\left(\beta,t+s+1\right)u_i\left[r_i\left(a_{t+s+1}\right)\right] \tag{B.6}$$

Therefore, putting together equations B.5 and B.6 we get

$$U_t^i = (1-\beta)\left[\frac{1}{t}+\beta\eta\left(\beta,t+1\right)\right]\left(\sum_{k=0}^{t} u_i\left[r_i\left(a_t\right)\right]\right) + \beta\left(1-\beta\right)\sum_{s=0}^{\infty}\beta^s\eta\left(\beta,t+s+1\right)u_i\left[r_i\left(a_{t+s+1}\right)\right]$$

and using the fact that $\eta\left(\beta,t+1\right) = \frac{1}{\beta}\left[\eta\left(\beta,t\right)-\frac{1}{t}\right] \Longleftrightarrow \frac{1}{t}+\beta\eta\left(\beta,t+1\right) = \eta\left(\beta,t\right)$, we then get that

$$U_t^i = (1-\beta)\,\eta\left(\beta,t\right)\sum_{k=0}^{t} u_i\left[r_i\left(a_t\right)\right] + \beta\left(1-\beta\right)\sum_{s=0}^{\infty}\beta^s\eta\left(\beta,t+s+1\right)u_i\left[r_i\left(a_{t+s+1}\right)\right]$$

as we wanted to show. □

Lemma B.4. *Define the repeated game* $\Gamma_t = \left\{\left\{A,\tilde{U}_t^i\right\}_{i=0}^{n}\right\}$ *which is identical to* Γ *in every aspect, only that sequences* $\mathbf{a} = \{a_s\}_{s=0}^{\infty}$ *are valuated as*

$$\tilde{U}_t^i\left(\mathbf{a}\right) = (1-\beta)\,u_i\left(a_0\right) + (1-\beta)\sum_{s=1}^{\infty}\left[\frac{\eta\left(\beta,t+1+s\right)}{\eta\left(\beta,t\right)}\right]\beta^s u_i\left[r_i\left(a_s\right)\right] \tag{B.7}$$

Then, under assumptions 1, 2 and 3, game Γ_t *is continuous at infinity (so the single deviation principle applies),* $\mathbf{SPE}\left(\Gamma_t\right) \neq \emptyset$*, the value set* $V(\Gamma_t)$ *is compact, and the following rule holds*

$$\bar{\boldsymbol{a}} = \{a_s\}_{s=0}^{\infty} \in \mathbf{O}\left(\Gamma_t\right) \Longleftrightarrow \exists \boldsymbol{a}^{t-1} : \left(\boldsymbol{a}^{t-1}\bar{\boldsymbol{a}}\right) \in \mathbf{O}\left(\Gamma\right) \tag{B.8}$$

Proof. Tychonoff's theorem implies that the set $A^{\infty}$is a compact set in the product topology. The functions $(1-\beta)\frac{\eta(\beta,t+s+1)}{\eta(\beta,t)}\beta^s u_i\left[r_i\left(.\right)\right]$is continuous, which implies that $\tilde{U}_t^i$ is continuous on A^{∞}in the product topology. Using these two facts, we replicate the proof of Proposition 2 in Abreu (1988) to show that $V(\Gamma_t)$ is compact. Condition B.8 follows from Proposition 5.3 and the fact that Γ_t is continuous at infinity. □

B.2. **Recursive Method for RAR payment.** The typical dynamic programming program involves solving

$$V\left(x_0\right)=\sup_{\{x_{t+1}\}_{t=0}^{\infty}}(1-\beta)\sum_{t=0}^{\infty}\beta^t F\left(x_t,x_{t+1}\right)$$

such that

$$\begin{cases} x_{t+1}\in G\left(x_t\right) & \forall t\in\mathbb{N} \\ x_0 & \text{given.}\end{cases}$$

The usual Bellman equation is

$$V\left(x\right)=\sup_{x'\in G(x)}(1-\beta)F\left(x,x'\right)+\beta V\left(x'\right).$$

If we allow for some random variable z and x' to be a function of z,

$$V\left(x,z\right)=\sup_{x'\in G(x,z)}(1-\beta)F\left(x,x',z\right)+\beta\mathbb{E}\left\{V\left(x',z'\right)\mid z\right\}.$$

Observe that, in contrast, when paying for a round at random, the problem is

$$V\left(x_0\right)=\sup_{\{x_{t+1}\}_{t=0}^{\infty}}(1-\beta)\sum_{t=0}^{\infty}\beta^t\eta\left(\beta,t+1\right)F\left(x_t,x_{t+1}\right)$$

such that

$$\begin{cases} x_{t+1}\in G\left(x_t\right) & \forall t\in\mathbb{N} \\ x_0 & \text{given.}\end{cases}$$

We will define $H\left(x_t,x_{t+1},\eta_t\right)=\eta_t F\left(x_t,x_{t+1}\right)$. In addition, given β, $\eta\left(\beta,t\right)$ is a strictly decreasing function of t. Therefore, let $T\left(\eta,\beta\right)$ be the inverse.[9] Using (B.1) and augmenting the state space with η, which has a known law of motion, yields

$$V\left(x_0,\eta_0\right)=\sup_{\{x_t,\eta_t\}_{t=0}^{\infty}}(1-\beta)\sum_{t=0}^{\infty}\beta^t H\left(x_t,x_{t+1},\eta_t\right)$$

such that

[9] $T\left(\eta,\beta\right)$ satisfies $\eta\left(\beta,T\left(\eta,\beta\right)\right)=\eta$.

$$\begin{cases} x_{t+1} \in G\left(x_t\right) & \forall t \in \mathbb{N} \\ \eta_{t+1} = \frac{1}{\beta}\left(\eta_t - \frac{1}{T(\eta_t,\beta)}\right) & \forall t \in \mathbb{N} \\ \eta_0 = \frac{1}{\beta}\ln\left(\frac{1}{1-\beta}\right) \text{ and } x_0 & \text{given.} \end{cases}$$

Therefore, the Bellman equation for this problem is simply

$$V\left(x,\eta\right) = \sup_{x' \in G(x)} \left(1-\beta\right)\eta F\left(x,x'\right) + \beta V\left(x',\eta'\right)$$

such that

$$\eta' = \frac{1}{\beta}\left(\eta - \frac{1}{T(\eta,\beta)}\right).$$

B.3. **Recursive Method for All Round Payment.** We want to get the euler equation in 6.3. We can characterize the optimal allocation by means of the following bellman equation:

$$V\left(a,S,y\right) = \max_{a'} \left(1-\beta\right) u\left(S+y+Ra-a'\right) + \beta V\left(a',S+y+Ra-a',y'\right)$$

where $S = \sum_{s=0}^{t} c_s$ is the sum of consumptions the agent would be paid if the game ended today. Assuming that the optimum is in the interior, we have

(B.9)

$$\left(1-\beta\right) u'\left(S+y+Ra-a'\right) = \beta\left\{V_a\left(a',S+y+Ra-a',y'\right) - V_S\left(a',S+y+Ra-a',y'\right)\right\}$$

and by the envelope conditions we have

$$V_a\left(a,S,y\right) = R\left(1-\beta\right) u'\left(S+y+Ra-a'\right) \tag{B.10}$$

$$V_S\left(a,S,y\right) = \left(1-\beta\right) u'\left(S+y+Ra-a'\right). \tag{B.11}$$

Using B.9 and substituting for c_t and c_{t+1}, we then get that

$$u'\left(c_t\right) = \beta\left(R-1\right) u'\left(c_{t+1}\right)$$

as we wanted to show.