Economic Theory
https://doi.org/10.1007/s00199-021-01409-x

RESEARCH ARTICLE

A note on payments in the lab for infinite horizon dynamic games with discounting

Arun Gautham Chandrasekhar[1,2] · Juan Pablo Xandri[3]

Received: 24 July 2019 / Accepted: 17 December 2021

Abstract
It is common for researchers studying infinite horizon dynamic games in a lab experiment to pay participants in a variety of ways, including but not limited to outcomes in all rounds or for a randomly chosen round. We argue that these payment schemes typically induce different preferences over outcomes than those of the target game, which in turn would typically implement different outcomes for a large class of solution concepts (e.g., subgame perfect equilibria, Markov equilibria, renegotiation-proof equilibria, rationalizability, and non-equilibrium behavior). For instance, paying subjects for all rounds generates strong incentives to behave differently in early periods as these returns are locked in. Relatedly, a compensation scheme that pays subjects for a randomly chosen round induces a time-dependent discounting function. Future periods are discounted more heavily than the discount rate in a way that can change the theoretical predictions both quantitatively and qualitatively. We rigorously characterize the mechanics of the problems induced by these payment methods, developing measures to describe the extent and shape of the distortions. Finally, we prove a uniqueness result: paying participants for the last (randomly occurring) round, is the unique scheme that robustly implements the predicted outcomes for any infinite horizon dynamic game with time separable utility, exponential discounting, and a payoff-invariant solution concept.

Keywords Payment in experiments · Experimental economics · Dynamic game experiments

JEL Classification C90 · C91 · C92

1 Introduction

A rapidly growing literature involves lab experiments designed to study behavior in infinite horizon dynamic games. The experimenter's goal, for the purposes of this

Both authors contributed equally to this work.

Extended author information available on the last page of the article

paper, is to test a model that predicts behavior given a physical environment. A *theoretical model* for the target game is a pair of (1) a model for agents' preferences and (2) a solution concept that predicts agents' behavior (e.g, subgame perfect equilibria, renegotiation-proof equilibria, rationalizability, Nash bargaining, though the behavior need not even be in equilibrium). To study the game, the researcher has to choose how to pay the subject, and the structure of the payment may influence the effective game that the subjects are playing.[1] In this paper, we consider implementing a broad class of infinite horizon dynamic games: those with models exhibiting discounting, time separability, and solution concepts that are payoff-invariant (that is, they depend only on agents preferences over game outcomes). This includes, but is not limited to, subgame perfect equilibria, perfect Bayesian equilibria, Nash bargaining and rationalizability. We demonstrate that the payment schemes used in the literature often implement different outcomes than the target game for a large class of solution concepts, by changing the underlying preferences of agents over game outcomes. We characterize the qualitative and quantitative differences and develop a payment scheme that robustly implements the target game for any infinite horizon dynamic game in the large class we consider, in the sense that it leaves preferences over outcomes identical to those of the target game. Finally, we demonstrate that the unique payment scheme that robustly implements the target game is paying participants for the last (randomly occurring) round.

Dynamic game experiments include analyses of growth models Lei and Noussair (2002), risk-sharing games (Charness and Genicot 2009; Chandrasekhar et al. 2011), common good extraction problems (Vespa 2013), investment choice and joint liability games (Fischer 2013), dynamic savings models (Brown et al. 2009), equilibrium refinements in dynamic games (Cooper et al. 1993), and of course classical infinitely repeated games (Dal Bo 2005). These studies vary not only by topic but also by the solution concepts considered. For instance, Vespa (2013) looks at Markov perfect equilibria in an experiment studying the common good extraction problem, Charness and Genicot (2009) focus on subgame perfection, Fischer (2013) is interested in perfect Bayesian equilibria, and Cooper et al. (1993) focus on forward induction.

Infinite horizon dynamic games are typically implemented in the lab using the *random termination method* and paying for all rounds or a random round. Typically, a participant plays a round of a game which then continues to the subsequent round with a given probability (Roth and Murnighan 1978). To incentivize behavior, the experimenter pays the participant as a function of the history of play. The central problem is that in the lab payments are made after the experiment and therefore not consumed between stages of the game (as they would be in the realm of the model). Experiments in the literature, following Murnighan and Roth (1983), usually pay subjects for *all rounds*. More recently Azrieli et al. (2018) systematically catalogue work in a collection of top journals and show that 56% pay for all rounds and 37.5% pay for one or several randomly chosen rounds.

Paying for all rounds is only valid when agents are assumed to be risk neutral (Murnighan and Roth 1983). While a section of the literature is interested in worlds respecting risk-neutrality, paying individuals for all rounds was (and often is) standard

[1] This has been the subject of a literature going back to Murnighan and Roth (1983).

even when the models being tested explicitly deviated from risk neutrality. This lead to a dissonance between the theoretical ambitions and experimental implementation of such work. Indeed, Azrieli et al. (2018) document that 48% of the papers they examine do not justify their payment scheme in the given experiment whatsoever.

There is a smaller section of the literature that implements random round payments and often this line of work is interested in within-round risk-averse behavior (e.g., insurance games, savings games, growth games) moved to paying individuals for a *randomly chosen round* (Davis and Holt 1993). An example of the argument offered is reproduced here:

> “As described in Charness and Genicot (2009), this [randomly chosen round] payment structure prevents individuals from self-insuring income risk across rounds. The utility maximization problem of the experiment corresponds to that of the theoretical model.” (Fischer, 2013, footnote 5).

Unfortunately, as our analysis shows, this is untrue.

Explicit risk aversion over rounds in a dynamic game experiment is important for three reasons. First, a growing literature is directly interested in risk-averse and risk-sharing behavior (Charness and Genicot 2009; Lei and Noussair 2002; Fischer 2013; Brown et al. 2009). Strict concavity of the utility of wealth is necessary for interior savings decisions as well as for any form of risk-sharing among two or more agents (Lei and Noussair 2002; Chandrasekhar et al. 2011; Brown et al. 2009). Second, a section of the literature is interested in high stakes games (Gneezy and Rustichini 2000). With high stakes comes risk aversion. This is especially true for lab-in-the-field work in the developing world, where payments are large relative to a day’s wage (Fischer 2013; Chandrasekhar et al. 2011). Third, recent work suggests that even in lower-stakes game we may observe risk aversion in the lab (Holt and Laury 2002; Harrison and Rutström 2008; Harrison et al. 2007). Given the possibility of curvature, it suggests that paying for all rounds or random rounds may not be *robust* payment schemes. We make this precise in our analysis, “ how” Showing, in an example arbitrarily small amounts of curvature may still generate deviations from the theoretical model.

This paper rigorously studies payment schemes in infinite horizon dynamic environments with time separable utility and exponential discounting. We formally introduce our environment in Sect. 2. We study a class of dynamic environments among a collection of agents in general terms. The researcher is interested in studying features of behavior in such a game and therefore constructs an environment in which the predictions of the game are borne out. The core concept therefore concerns whether the experimental environment, inclusive of the payment scheme for the subjects, implements the game of interest. We will say the scheme *implements the game* if the preferences over outcomes for all agents along all histories coincide. Note that this includes but is not limited to equilibrium behavior. It ensures that along all paths, the behavior in the experiment exactly mimics the behavior in the target game for almost any rational solution concept.

In Sect. 3, we introduce and study the scheme wherein individuals are paid for the *last round* in a random-termination dynamic game. It implements all such models.[2] The reason why this works is due to the (standard) observation that when we have exponential discounting, myopia with respect to the future is isomorphic to random termination of the game with some probability.[3] Further, we show that in fact last round payment is the unique payment scheme that implements the game robustly. What this means is that for a given environment, one can find some agent and some history where for some preference (e.g., small amounts of curvature in some cases) the scheme fails to implement the target game. That is, the preference ordering for the agent in that situation is muddled and does not reflect that within the target game.

Next, in Sect. 4, we provide a rigorous analysis of two payment schemes used in the literature – either payment for a randomly chosen round or all rounds – and show how these may induce games different from the target game. We are able to measure the extent and shape of the distortions. For instance, paying subjects for a randomly chosen round induces a time-dependent discounting function. Future periods are discounted more heavily than the discount rate in a way that can change the theoretical predictions both quantitatively and qualitatively. We compute the exact time-varying discount function that is induced by the payment scheme, characterize its properties, and study the behavior in the induced game. Our results demonstrate that distortions are sizable and can be quantitatively and qualitatively important. For any discount rate and any dynamic game, consider the "virtual" net present value (NPV) of a constant stream of consumption. In the first round, under payment for a randomly chosen round, would be at the most half of the actual NPV under the true model. In particular, if the discount rate is 0.95, the "virtual" NPV would be less than 30% of the value under the theoretical model. Moreover, paying in all rounds creates distortions when we allow for arbitrarily small amounts of curvature in the utility for wealth. In addition to generating distortions, these payment schemes are not robust. Asymptotically, subjects should care less and less about their decisions in a precise sense. The maximal difference in utility due to any two future sequences of actions given any history should go to zero add the number of rounds placed increases. This means that small perturbations to the system can lead to drastically different observed outcomes.

We illustrate the pathologies in Sect. 5, presenting our results in the context of a single-person savings and borrowing model. Whereas an individual should be consuming a constant or growing amount under the theoretical model, randomly chosen round and all round payment can generate a decreasing consumption sequence. This can lead to qualitatively misleading results which could, for instance, be interpreted

---

[2] Concurrently and independently of us Sherstyuk et al. (2013) also introduce last round payment in infinite horizon games in a repeated prisoner's dilemma experiment.

[3] However, to be able to implement models with hyperbolic discounting with this payment method, a researcher should believe that this sort of behavioral bias is isomorphic with the inability to calculate probabilities, which is likely to be unreasonable. This comes from the fact that if we try to map $(1-\beta)\,u\,(c_t) + \beta\gamma W\,(t+1)$ (where W is the continuation value and $\gamma \in (0,1)$) to random termination it must be the case that $\Pr(\text{game ends}) + \Pr(\text{game continues}) = (1-\beta) + \beta\gamma < 1$. Equivalently, the agent does not know how to calculate probabilities.

as “learning to save” as the game progresses; however, no learning is occurring and this is a product of incentives.

Our results suggest that for a broad class of dynamic games – infinite horizon with time separability, exponential discounting, and payoff-invariant solution concepts – experimentalists can use last round payment under random termination as a robust, theoretically-justified implementation scheme. Irrespective of the solution or extent of risk aversion, the scheme will implement the target game under standard assumptions. Additionally, our results allow a researcher to correct her theoretical predictions if she was interested in estimating a structural model from the lab experiment data utilizing an alternative payment scheme. 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 estimation.

Our paper is contributes to a recent push in the literature to study the effect of incentives in experiments. The papers closest to ours are complementary investigations by Charness and Genicot (2009) and Sherstyuk et al. (2013). Charness and Genicot (2009) conduct an infinite horizon risk-sharing experiment. Noting that all round payments will not work for their purposes, they pay for a randomly chosen round. They look at the discount factor induced by random round payment and show that the incentive compatibility constraints of the induced game under random round payment converges to the target games in later rounds. Below we show that this is insufficient to argue that later rounds of the experiment should represent equilibrium behavior in the target game (and generally will fail to do so). Sherstyuk et al. (2013), concurrently and independently of us, look at last round payment in addition to all round and randomly chosen round payment in the context of infinitely repeated games in risk neutral settings. In an experiment implementing infinitely repeated Prisoner’s Dilemma, they confirm that in risk-neutral settings paying for a randomly chosen round will not implement the target game whereas both all round and last round payment will. They also confirm Charness and Genicot (2009)’s characterization of the limiting discount factor, though they argue this induces time inconsistency. As a byproduct of our general analysis, it is easy to see that the problem is in fact time consistent, with a time varying discounting function that we exactly characterize. Most recently, Azrieli et al. (2018) both survey the techniques used in the literature and also look at when specific schemes such as all round or random round payment can be justified.

The work in this paper also fits in with a broader experimental methods literature on payment techniques For instance, many experiments conduct multiple sessions. In our context, this means each subject plays multiple, independent, dynamic games. Azrieli et al. (2012) look at how payments should be structured to decouple interlinkages across several dynamic games, though does not investigate how to implement the dynamic game itself. In their setup, their concept of implementability is akin to our definition of robust implementation, and they obtain that in fact, a mechanism paying the agent for a randomly chosen task. This is compatible with our results, since we focus on the implementability conditions for a single game (a dynamic one). Therefore, if we were testing a number of dynamic games in the lab, we would pay for a randomly chosen game, but the payment for that game should implement the same preferences, which is achieved by paying for the last round of the randomly chosen game. Fréchette and Yuksel (2013) investigate how subjects treat random termination, since the Murnighan and Roth (1983) procedure links the expected number of rounds with the discount factor. They study three alternatives, including Cabral et al. (2011)'s approach of fixing a set number of rounds and then employing random termination. Our paper takes the standard Roth and Murnighan (1978) random termination as the benchmark; studying what payment schemes are optimal under specific behavioral departures from the classical model requires adding these assumptions to the model.

Importantly, our analysis highlights a robustness of last round payment, even in risk neutral environments. Recall that it works under any solution concept that is payoff-invariant, meaning it depends only on the conditional expected utility of agents, nesting SPE, Nash bargaining, rationalizability, etc. Thus, from a theoretical perspective we view it as weakly dominant to use as a payment scheme, particularly in light of the uniqueness result. Finally, if researchers are reluctant to implement this payment scheme for behavioral reasons (outside the current model), we note that any such behavioral frictions ought to be modeled in so that the researcher can understand both the likely distortions and the optimal implementing payment mechanism under the augmented model. While this is beyond the scope of our paper, the approach would be straightforward using our techniques.

2 Framework

As noted before our treatment applies to multi-stage games with payoff-invariant solution concepts; i.e, solution concepts that are only a function of the preferences of the environment and preferences of the agents. This covers a wide array of experimental environments with dynamic games: repeated games, bargaining games, etc. Additionally, this covers a large class of solution concepts for these games: SPE, Markov equilibria, renegotiation proof equilibria, PBE, sequential rationality, forward induction refinements (intuitive criterion, stability), etc. We do so by focusing on choosing a payment scheme that implements the same preferences, so any payoff-invariant solution concept would coincide with that of the target game. We consider the general analysis in Appendix A where we extend our results to a general class of multi-stage games with time separable utility and (possibly) history dependent discounting.

2.1 Setup and notation

The researcher is interested in conducting an experiment to test behavior in a dynamic game Γ. The game form (i.e, the physical environment) is specified as

$$E := \left\{ \left\{ A_i, r_i : X \times \prod_{i=1}^{n} A_i \to \mathbb{R}_{\geq 0} \right\}_{i=1}^{n}, \{X, G : A \times X \to X\}, x_0 \in X,\ \beta \in (0,1) \right\}$$

where A_i is the stage action space for agent i, $A = \prod_{i=1}^{n} A_i$ the set of action profiles, and $r_i(\cdot)$ are the monetary reward functions for each agent. The set X of states evolves according to $x_{t+1} = G(a_t, x_t)$, and x_0 is the first (given) realization of the state. The special case where $G(a, x) = \{x_0\}$ would correspond to a standard repeated game. For simplicity, we will assume that A and X are compact topological spaces.

Let $\mathcal{H} \subseteq \cup_{t=0}^{\infty} (A \times X)^t$ be the set of all possible histories; i.e, those that satisfy that $x_s = G(a_{s-1}, x_{s-1})$ for all $s \leq t$. A round t outcome is written as $h_t = (a_t, x_t)$. A round t history is composed by the outcomes up to $t-1$; i.e $h^t = (a_s, x_s)_{s \leq t-1}$ such that $x_s = G(a_{s-1}, x_{s-1})$. 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 $h^t = (h_0, h_1, ..., h_{t-1})$ where $h_k = (a_k, x_k)$, an action $a_{i,t} = \sigma_i\left(h^t\right)$.

We will also denote $\sigma_i | h^t$ to be the *continuation strategy* induced by history h^t. Given a strategy profile $\sigma = (\sigma_1, ..., \sigma_n)$, the *outcome of* σ is a sequence $\mathbf{o}(\sigma) \in (A \times X)^{\infty}$ of actions and states prescribed by σ. Let Σ be the set of strategy profiles for environment E.

To define a game Γ based on environment E, we need to specify continuation preferences of agents over *feasible outcome sequences* $\mathbf{o} = \left\{\hat{a}_s, \hat{x}_s\right\}_{s=0}^{\infty} \in (A \times X)^{\infty}$ such that $\hat{x}_{s+1} = G\left(\hat{a}_s, \hat{x}_s\right)$ for all s. Given two histories $h^t = (h_k)_{k \leq t-1}$ and $h^s = (\hat{h}_j)_{j \leq s-1}$ where $\hat{h}_0 = h_t$, we write the *concatenation* of them as a new history $h^{t+s} := (h^t \hat{h}^s)$. We say that a feasible outcome $\mathbf{o} = \left(\hat{a}_s, \hat{x}_s\right)_{s=0}^{\infty}$ *is consistent with* $h^t = (a_k, x_k)_{k \leq t-1}$ if $\left(\hat{a}_0, \hat{x}_0\right) = (a_t, x_t)$; i.e, it starts with the last state in h^t, and is a feasible sequence given the law of motion for the state x. Moreover, we write $h^t_{\mathbf{o}} = \left(\hat{a}_k, \hat{x}_s\right)_{k \leq t-1}$ as the $t-$history consistent with outcome $\mathbf{o}$.

See that E does not specify preferences for agents, and is therefore not a complete description of a game. Let $\mathbf{u} := (u_i : \mathbb{R} \to \mathbb{R})_{i=1}^{i=n}$ denote a particular profile of monetary utility functions, and define the dynamic game $\Gamma_{\mathbf{u}} := \left\{A_i, X, G : A \times X \to X, \hat{u}_i : \mathbb{R}_+ \to \mathbb{R}, \beta\right\}$ where $\hat{u}_i(a, x) := (u_i \circ r_i)(a, x)$. Given utility functions $u_i : \mathbb{R}_+ \to \mathbb{R}$, the continuation utility of an outcome $\mathbf{o}$ consistent with $h^t \in \mathcal{H}$ is given by the time separable utility function:

$$U_i\left(\mathbf{o} \mid h^t\right) = (1-\beta) \sum_{s=0}^{\infty} \beta^t u_i\left[r_i(a_{t+s}, x_{t+s})\right] \tag{1}$$

where $\mathbf{o} \mid h^t$ is the tail (from period t on) of the sequence $\hat{\mathbf{o}} = \left(h^t \mathbf{o}\right)$ (i.e, $\mathbf{o}$ is the continuation outcome). The results we obtain for these environments are easily generalizable for the case where x_t includes a random markov process. Formally, when $x_t = (z_t, y_t)$, where $z_t = Z(a_{t-1}, x_{t-1})$ is a deterministic state, and y_t is a markov process with conditional distribution $y_t \sim F(y \mid x_t)$. In this case, we write $u_i\left[r_i(a_{t+s}, x_{t+s})\right] = \int \hat{u}_i\left[r_i(a_{t+s}, z_{t+s}, y_{t+s})\right] dF(y \mid x_t)$ for the underlying monetary utility function $\hat{u}_i(\cdot)$ [4]

2.2 Payment schemes and implementation

The researcher wants to test $\Gamma = \Gamma_{\mathbf{u}}$ for a particular class of utility functions $\mathbf{u}$, which has an infinite horizon and in which agents have exponential discounting. As such,

[4] This corresponds to the case with public random variables. However, is also easy to provide an extension where agents also have sequential private information.

she has to design an alternative repeated game $\widehat{\Gamma}$ with finite (but random) horizon with the same strategy space, to test the predictions of game Γ. It is a well-known fact (following Roth and Murnighan (1978)) that the infinite horizon and exponential discounting nature of Γ can be replicated by a game that ends in T periods, where T is a geometric random variable with probability of termination β, whose realization is not known to the players; $\Pr(T = t) = (1-\beta)\,\beta^t$. We write $T \sim \text{Geom}(\beta)$.

To test the predictions of game Γ, the researcher chooses a *payment scheme* $\mathbf{R} = \left\{R(h^t)\right\}_{h^t \in \mathcal{H}}$ that specifies, for each history h^t, the payments $r = (r_i)_{i\in[n]}$ in terms of a distribution $R(h^t) \in \Delta(\mathbb{R}^n)$ of monetary rewards *at the end of the game*; i.e, if the last period agents play is $T = t$. For example, she might pay choose an *all rounds payment scheme*: i.e, $R(h^t) = \prod_{i=1}^n R_i(h^t)$ with $R_i\left(h^t\right) := \left\{\sum_{s\leq t-1} r_i\left(a_s, x_s\right) \text{ with prob. } 1\right\}$ which we denote $\mathbf{R}^{all}$. Another salient example is to choose the following payment scheme:

$$R_i(h^t) = r_i(a_s, x_s) \text{ with prob. } \frac{1}{t+1} \text{ for all } s \leq t-1.$$

We will call this the *randomly chosen round* payment scheme (where the choice is uniform at random) and we will denote it $\mathbf{R}^{rcr}$. For a given history h^s, we write $R_i(h^t \mid h^s)$ as the distribution for the history the payment scheme induces a new dynamic game, which we will call $\widehat{\Gamma}(\mathbf{R})$. At any history h^t, the individual evaluates future sequences according to:

$$\widehat{U}_i^{\mathbf{R}}\left(\mathbf{o} \mid h^t\right) := E_T\left\{E_r^{\mathbf{R}}\left[u_i\left(r_i\right) \mid h^T h^t_{\mathbf{o}}\right]\right\} \tag{2}$$

where $T \sim \text{Geom}(\beta)$ and $E_r^{\mathbf{R}}\left(\cdot \mid h^t\right)$ is the expectation (over r) using the probability distribution $\mathbf{R}(h^t)$. That is, the outcome $\hat{\mathbf{o}} = (h^t\mathbf{o})$ generates partial finite histories $h^{t+T} = h^T h^t_{\mathbf{o}}$. If the researcher wants to test the predictions of game Γ, the payment scheme used *must not change the preferences of the original game* Γ. This is formalized in the following definition.

Definition 1 (Implementation) We say a *payment scheme* $\mathbf{R} = \left\{R(h^t)\right\}_{h^t \in \mathcal{H}}$ implements $\Gamma = \Gamma_{\mathbf{u}}$ if the preferences over action sequences coincide; i.e, for all feasible $\mathbf{o}, \mathbf{o}' \in (A \times X)^\infty$ and all $h^t \in \mathcal{H}$ we have that

$$U_i\left(\mathbf{o} \mid h^t\right) \geq U_i\left(\mathbf{o}' \mid h^t\right) \text{ if and only if } \hat{U}_i^{\mathbf{R}}\left(\mathbf{o} \mid h^t\right) \geq \hat{U}_i^{\mathbf{R}}\left(\mathbf{o}' \mid h^t\right)$$

for all $i = 1, \ldots, n$.

Definition 2 (Strong Implementation) We say a *payment scheme* $\mathbf{R} = \left\{R(h^t)\right\}_{h^t \in \mathcal{H}}$ strongly implements an environment E if it implements the game $\Gamma = \Gamma_{\mathbf{u}}$ for all utility function profiles $\mathbf{u} = (u_i(\cdot))_{i=1}^n$.

We refer to such schemes as *implementing and strong implementing schemes* respectively. This definition of implementation is useful in that it does not make any assumptions about equilibrium selection, or common knowledge assumptions. It also implies that any solution concept that is based on the agents preferences will be the same in both games. We refer to this solution concepts as payoff invariant solutions, which we study in Appendix A. Strong implementation is useful for studies that also aim to identify (or to test) the preferences of agents in the stage game.

Given a payment scheme **R**, we are interested in studying whether it implements Γ. From Definition 1 we see that the time separability assumption implies that *any* implementing scheme must be "memoryless" in the sense that continuation preferences after history h^t, should be the independent of past histories $h^s \prec h^t$. If the the payment scheme has history dependence, we might suspect that it is not an implementing scheme.

We will show that both paying individuals for a randomly chosen round and paying individuals for all rounds (with any amount of curvature of the utility function) will typically violate this property. This is because when we pay for a randomly chosen round, the incentives faced by the agent will depend explicitly on the round number t and when we pay for all rounds, the payment scheme explicitly depends on the entire history. However, in the next section we show the existence of an implementing payment scheme that essentially works for any game.

2.3 Payoff-invariant solution concepts.

By focusing on maintaining preferences of agents maintained (no matter the per-period utility functions) we can also respect, in the induced lab game, the same solution concept that the experimenter aims to test. In general, a solution concept $\mathcal{S}(\Gamma)$ is a function that, for each game $\Gamma = \Gamma_{\mathbf{u}}$ based on environment E, it predicts that observed outcomes will be a subset $\mathcal{S}(\Gamma) \subseteq \mathcal{H}$. We say that $\mathcal{S}$ is payoff-invariant if, whenever we have games Γ, Γ' based on E with the same preferences over outcomes (i.e, they are *payoff-equivalent*)[5], then we should also have $\mathcal{S}(\Gamma) = \mathcal{S}(\Gamma')$. Typical examples include Nash and Subgame Perfect equilibria, Markov equilibria and renegotiation-proof refinements, Pareto optimal outcomes, cooperative bargaining solutions, and so on.

Our implementation definition guarantees that if **R** implements Γ, then $\hat{\Gamma}(\mathbf{R})$ is payoff equivalent to Γ, and hence any payoff-invariant solution concept $\mathcal{S}$ would give the same prediction $\mathcal{S}\left(\hat{\Gamma}(\mathbf{R})\right) = \mathcal{S}(\Gamma)$.

3 Last round payment

3.1 Implementation

Define the scheme $\mathbf{R}^{last} = (\pi_{i=1}^{n} R_i(h^t)_{h^t \in H}$ where

$$R_i^{last}(h^t) := \{r_i(a_{t-1}, x_{t-1}) \text{ with prob. } 1\}$$

[5] For game $\Gamma_{\mathbf{u}}$, define $U_i^{\mathbf{u}}(\mathbf{o} \mid h)$ as the continuation utility of outcome $\mathbf{o}$ after history h (as in equation 1). We say that $\Gamma = \Gamma_{\mathbf{u}}$ and $\Gamma' = \Gamma_{\mathbf{u}'}$ are *payoff equivalent* if, for all i, $h \in \mathcal{H}$ and outcomes $\mathbf{o}, \mathbf{o}'$, we have $U_i^{\mathbf{u}}(\mathbf{o} \mid h) \geq U_i^{\mathbf{u}}(\mathbf{o}' \mid h) \iff U_i^{\mathbf{u}'}(\mathbf{o} \mid h) \geq U_i^{\mathbf{u}'}(\mathbf{o}' \mid h)$

and $T \sim \text{Geom}(\beta)$. Here we pay agents for only the last period of the game. Having arrived at a history h^t, the payment scheme prescribes that with probability $1-\beta$ the game ends and agents get as final reward $r_i(a_t, x_t)$. 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. As this payment scheme does not depend on the history h^t, the implementability test is satisfied.

Proposition 1 *Payment scheme* $\mathbf{R}^{last}$ *strongly implements environment* E

Proof Fix a profile of utility functions $(u_i)_{i \in I}$. For any history h^t and any outcome $\mathbf{o} \mid h^t$ we have that

$$\widehat{U}_i^{\mathbf{R}^{last}}\left(\mathbf{o} \mid h^t\right) = E_T\left\{E_r^{\mathbf{R}^{last}}\left[u_i\left(r_i\right) \mid h^T h_{\mathbf{o}}^t\right]\right\} = E_T\left\{u_i\left[r_i\left(a_{t+T}, x_{t+T}\right)\right]\right\}$$

and using the fact that $T \sim \text{Geom}(\beta)$ we then have that

$$\widehat{U}_i^{\mathbf{R}^{last}}\left(\mathbf{o} \mid h^t\right) = \sum_{s=0}^{\infty}(1-\beta)\,\beta^t u_i\left[r_i\left(a_{t+s}, x_{t+s}\right)\right] = U_i\left(\mathbf{o} \mid h^t\right)$$

Since this property holds for any profile of utility functions, we show that $\mathbf{R}^{last}$ strongly implements E. □

This argument is easily generalizable for *any* multistage game with observable actions, time separable utility and common discount factors, which we do in Appendix A.

An important extension is for environments with non-exponential discounting; e.g. $U_i(\mathbf{o}) = \sum_{t=0}^{\infty} b_t u_i\left(r_{i,t}\right)$ where $b_t \geq 0$ and $\sum_t b_t = 1$. In this case, last round payments also implement this environment, but with a different random process for the number of rounds chosen. Namely, a random length is drawn from the distribution $\Pr(T = t) = b_t$, but is only revealed, at every stage, whether the game ends or at least goes for another round. Namely, if we have reached round t, the probability that the game ends in this round is $1 - \beta_t := \Pr(T = t \mid T \geq t)$ which is here calculated as $b_t / \left(1 - \sum_{s<t} b_s\right)$, and it continues at least one period with probability β_t.[6]

When agents are risk neutral, we can implement the game (but obviously not strongly) more easily: write the discount factor as $b_t = \beta^t Q_t$ for some chosen $\beta \in (0, 1)$, where the term Q_t would be interpreted as a "reward discount". Then, instead of paying $r_{i,t}$ on the last round, we pay $\hat{r}_{i,t} = Q_t r_{i,t}$. It is trivial to check that this payment scheme implements the game with risk neutral agents.

3.2 Uniqueness

While we know that the payment scheme $\mathbf{R}^{last}$ is an implementing scheme, we show below that it is not the unique implementing scheme for all possible models. In particular, we will show that paying agents for all rounds, if agents had linear utility, it would

[6] Again, we note that 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.

also replicate the same preference ordering over continuation outcomes. However, we then will note that this is not robust; any deviation from linearity can generate very different preferences from the theoretical ones. So, when one wants their payment scheme to be robust to misspecification (or small deviations) of the utility functions over monetary rewards, the relevant implementation concept should be that of *strong implementation*. We now show a uniqueness result: in fact, $\mathbf{R}^{last}$ is the only strongly implementing scheme: for any other payment scheme $\mathbf{R}$, we can find a pair of outcome sequences, an agent i with a strictly increasing and smooth (i.e, infinitely derivable) utility function $u_i(\cdot)$ such that the induced preferences under $\mathbf{R}$ do not coincide with those of the theoretical model.

We will only consider schemes $\mathbf{R}$ where all lotteries over monetary rewards have finite support (which is reasonable in practical applications), A payment scheme $\mathbf{R}$ *disagrees* with $\mathbf{R}^{last}$ if we can find a pair of histories $h^t = (h^{t-1}h_{t-1})$ and $\hat{h}^t = (\hat{h}^{t-1}\hat{h}_t)$ with $h^{t-1} \neq \hat{h}^{t-1}$ and $h_{t-1} = \hat{h}_{t-1}$ (i.e, they have the same length, and coincide in the last period) such that $R_i(h^t) \neq R_i(\hat{h}^t)$ for some player i.

We say that a payment scheme $\mathbf{R}$ satisfies *quasi-separability* if, for a pair of histories $h^t, \hat{h}^t$, with $h_{t-1} = \hat{h}_{t-1} = (a_{t-1}, x_{t-1})$, and a strictly increasing utility function $u : \mathbb{R} \to \mathbb{R}$, $E_r^{\mathbf{R}}\left\{u(r) \mid h^t\right\} > E_r^{\mathbf{R}}\left\{u(r) \mid \hat{h}^t\right\}$ implies that $E_r^{\mathbf{R}}\left\{u(r) \mid h^t h^s\right\} \geq E_r^{\mathbf{R}}\left\{u(r) \mid \hat{h}^t h^s\right\}$ for any $h^s = (\hat{a}_k, \hat{x}_k)_{k \leq s-1}$ with $(\hat{a}_0, \hat{x}_0) = (a_{t-1}, x_{t-1})$. Note that $\mathbf{R}^{rar}$, $\mathbf{R}^{all}$ and $\mathbf{R}^{last}$ satisfy this property, as does the theoretical environment E.

We show that any quasi-separable payment scheme $\mathbf{R}$ that disagrees with $\mathbf{R}^{last}$ cannot strongly implement E. For this, we use a key property of time separable preferences, that states that if two histories $h^t, \hat{h}^t$ have the same final period outcome, then we must have that $U_i\left(\mathbf{o} \mid h^t\right) = U_i\left(\mathbf{o} \mid \hat{h}^t\right)$ for any outcome $\mathbf{o} = \{a_s, x_s\}_{s=0}^{\infty}$ consistent with h^t, and therefore also consistent with $\hat{h}^t$ (since $h_{t-1} = \hat{h}_{t-1}$). However, if $\mathbf{R}$ disagrees with $\mathbf{R}^{last}$, we can then find a utility function for which the expected utility at h^t is strictly greater (or smaller) that the expected utility at $\hat{h}^t$. Then, if $\mathbf{R}$ is also quasi-separable, any continuation outcome will be better appended to h^t than to $\hat{h}^t$ (or vice versa).

Proposition 2 *Let* $\mathbf{R}$ *be a finite support payment scheme that disagrees with* $\mathbf{R}^{last}$ *for agent i at histories* $h^t, \hat{h}^t$. *Then, we can find a smooth, strictly increasing, and eventually concave*[7] *utility function* $u_i : \mathbb{R}_+ \to \mathbb{R}$ *such that* $E^{\mathbf{R}}\left\{u_i(r) \mid h^t\right\} \neq E^{\mathbf{R}}\left\{u_i(r) \mid \hat{h}^t\right\}$. *If* $\mathbf{R}$ *is quasi-separable and* $E^{\mathbf{R}}\left\{u_i(r) \mid h^t\right\} > E^{\mathbf{R}}\left\{u_i(r) \mid \hat{h}^t\right\}$, *then* $U_i^{\mathbf{R}}(\mathbf{o} \mid h^t) > U_i^{\mathbf{R}}(\mathbf{o} \mid \hat{h}^t)$ *for any consistent outcome* $\mathbf{o} = (\hat{a}_s, \hat{x}_s)_{s=0}^{\infty}$ *with* $(\hat{a}_0, \hat{x}_0) = (a_{t-1}, x_{t-1})$.

Proof See Appendix B. □

We obtain from Proposition 2 the following Corollary

Corollary 3 *Take a finite support payment scheme* $\mathbf{R}$ *that is quasi-separable and disagrees with* $\mathbf{R}^{last}$. *Then, there exist a strictly increasing, smooth, and eventually concave utility function* $u_i : \mathbb{R} \to \mathbb{R}$ *such that* $\mathbf{R}$ *does not implement* $\Gamma = \Gamma_u$.

[7] That is, there exists $\hat{r}$ such that $u''(r) < 0$ for any $r > \hat{r}$, and $u''(r) > 0$ For any $r < \hat{r}$.

Proof See Appendix B. □

This corollary implies that $\mathbf{R}^{last}$ is the only payment scheme that strongly implements E. Of course, one might not be really interested in the payment scheme being robust for all possible utility functions, but for a strict subset of feasible utility functions $\mathcal{U} \subseteq \mathbb{R}^{\mathbb{R}_+}$ (e.g. concave functions, linear functions). Our proof finds a utility function $u_i(c)$ that is strictly increasing and smooth, exploiting the richness of the set of such functions to find an example that exactly makes the comparison strict (e.g. u_i is not everywhere concave). It would be interesting to study whether other payment schemes also strongly implement the environment, when the set of possible utility functions is further restricted.

4 Payment schemes in the literature

We now rigorously study the two payment schemes used in the vast majority of experiments catalogued in Azrieli et al. (2018).[8] First, in Sect. 4.1 we examine payment for a randomly chosen round and in Sect. 4.2 we look at the all round payment scheme.

4.1 Payment for a randomly chosen round

We argue that paying for a randomly chosen round does not generally implement Γ. Moreover, we characterize the behavior induced by this payment scheme. Individuals discount the future too much in any period and become asymptotically indifferent between their choices. We develop a formal measure of the distortion to be able to quantify these biases.

We begin by defining the functions

$$\eta(\beta, t) := \sum_{k=0}^{\infty} \frac{\beta^k}{t+k} \text{ and } \eta_\beta(\beta, t) := \frac{\partial \eta(\beta, t)}{\partial \beta}. \tag{3}$$

We catalog their properties in Appendix C. These are modified discount factors and, as we will show in Lemma 12, allow us to write down exactly how randomly chosen round payment can be represented as discounting with $\eta(\beta, t)$. Specifically, utility of outcome $\mathbf{a}$ from time $t = 0$ onward can be written as

$$\widehat{U}_i^{\mathbf{R}^{rcr}}(\mathbf{o}) = (1-\beta) \sum_{t=0}^{\infty} \eta(\beta, t) \beta^t u_i\left(r_{i,t}\right) \tag{4}$$

where $r_{i,t} = r_i(a_t, x_t)$. At any history h^t

[8] Related payment schemes such as in Cabral et al. (2011) can be analyzed using this approach.

$$\widehat{U}_i^{\mathbf{R}^{rcr}}\left(\mathbf{o}\mid h^t\right) = (1-\beta)\,\eta\,(\beta,t)\sum_{k=0}^{t} u_i\,(r_{i,k}) \\ +\beta\left\{(1-\beta)\sum_{s=0}^{\infty}\beta^s\eta\,(\beta,t+1+s)\,u_i\,(r_{i,t+s+1})\right\}. \quad (5)$$

From (4) we see that the implied discount factor is $\beta^t\eta\,(\beta,t)$ instead of just β^t. This means that individuals in game $\widehat{\Gamma}\,(\mathbf{R}^{rcr})$ discount future flows of utility too rapidly in the induced game relative to Γ. Moreover, we can see this from the fact that $\lim_{t\to\infty}\eta\,(\beta,t)=0$ (see, Lemma 10) that

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

As a practical matter, we should expect agents to behave much more impatiently than in the target model. 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 $(1-\beta)\,u_{i,1}$ utils at $t=1$, but that this also affects the expected utility at time $t=2$ by $(1-\beta)\,\beta\frac{1}{2}u_{i,2}$, at time $t=3$ by $(1-\beta)\,\beta^2\frac{1}{3}u_{i,3}$, 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. More explicitly, given a history h^t and two outcomes $\mathbf{o}$ and $\hat{\mathbf{o}}$, the agent compares utilities in the target theoretical game as

$$U_i\left(\hat{\mathbf{o}}\mid h^t\right)\geq U_i\left(\mathbf{o}\mid h^t\right) \\ \iff (1-\beta)\,u_i\,(\hat{r}_i)+\beta\left[(1-\beta)\sum_{s=0}^{\infty}\beta^s u_i\,(\hat{r}_{i,t+s})\right] \\ \geq (1-\beta)\,u_i\,(r_i)+\beta\left[(1-\beta)\sum_{s=0}^{\infty}\beta^s u_i\,(r_{i,t+s})\right] \quad (6)$$

where $\hat{r}_t=r_i\left(\hat{a}_t,\hat{x}_t\right)$ and $\hat{x}_{t+1}=G\left(\hat{a}_t,\hat{x}_t\right)$

However, when we try to implement it with the random chosen round payment method, the agent compares utilities according to a different utility function:

$$\widehat{U}_i^{\mathbf{R}^{rcr}}\left(\hat{\mathbf{o}}\mid h^t\right)\geq \widehat{U}_i^{\mathbf{R}^{rcr}}\left(\mathbf{o}\mid h^t\right) \\ \iff (1-\beta)\,u_i\,(\hat{r}_{i,t})+\beta\left[(1-\beta)\sum_{s=0}^{\infty}\frac{\eta\,(\beta,t+1+s)}{\eta\,(\beta,t)}\beta^s u_i\,(\hat{r}_{i,t+s})\right] \\ \geq (1-\beta)\,u_i\,(r_{i,t})+\beta\left[(1-\beta)\sum_{s=0}^{\infty}\frac{\eta\,(\beta,t+1+s)}{\eta\,(\beta,t)}\beta^s u_i\,(r_{i,t+s})\right]. \quad (7)$$

Condition (7) is nearly identical to the incentive condition (6) of the theoretical game Γ. The crucial distinction is 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,

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

Consequently a participant is more impatient in $\widehat{\Gamma}(\mathbf{R}^{rcr})$ than in Γ. In addition, punishments after deviations (i.e, the worst equilibrium payoffs for each agent) 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.

It turns out that in the long run (for t large enough), the incentive condition in sequences in Γ and $\widehat{\Gamma}(\mathbf{R}^{rcr})$ are arbitrarily close. Charness and Genicot (2009) document this phenomenon in a model of risk sharing with limited commitment. We confirm this generally in Lemma 11, showing that for any $s \in \mathbb{N}$

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

This means as players keep on playing, the incentive compatibility (IC) constraints of the actual game and the induced game are not very different.

Charness and Genicot (2009) use this to argue that $\mathbf{R}^{rcr}$ almost implements Γ in the sense that later rounds are likely to match the theoretical predictions. We caution that this argument does not follow, for several reasons. First, the speed of convergence of the incentive compatibility constraints is slow, as discussed below. Second, 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 of fixed income streams, which we show below converges at a much slower rate. Third, we prove that the participants will exhibit asymptotic indifference: even though the IC constraints of the target 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.

4.1.1 Asymptotic indifference

To study the asymptotic indifference, we begin by defining the contribution of an action to a payoff.

Definition 3 (Contribution) Let $K : (A \times X)^{\infty} \to \mathbb{R}$ be some function that can be written as

$$K(\mathbf{o}) = \sum_{s=0}^{\infty} F(s, a_s, x_s)$$

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

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

i.e, it is the increment in K caused by what happens at period t. Likewise, let $I \subset \mathbb{N}$ index set. We define the *contribution of* $\{a_k, x_k\}_{k\in I}$ as

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

Example 1 Take $U_{i,t} = (1-\beta)\sum_{s=t}^{\infty} \beta^{s-t} u_i(r_{i,t})$ as the time t utility for agent i in game Γ at all subgames that start at date t. Then

$$\mathcal{C}\left(U_{i,t} \mid t\right)(\mathbf{o}) = (1-\beta)\, u_i(r_{i,t})$$

and the contribution of all future utility flows is

$$\mathcal{C}\left(U_{i,t} \mid s > t\right)(\mathbf{a}) = (1-\beta) \sum_{s=t+1}^{\infty} \beta^{s-t} u_i(r_{i,s}).$$

For the particular case of a stationary path (i.e, $u_i(r_i) = \hat{u}_i$ for all s) then we can simplify the above as

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

Note that both are time independent and non-negligible. We also define the relative contribution $\mathcal{RC}_t$ as the contribution of present and future utilities, relative to the contribution of both past and present actions:

$$\mathcal{RC}_t\,(\mathbf{o}) = \frac{\mathcal{C}\left(U_t^i \mid s \geq t\right)(\mathbf{o})}{\mathcal{C}\left(U_t^i \mid s < t\right)(\mathbf{o}) + \mathcal{C}\left(U_t^i \mid s \geq t\right)(\mathbf{o})}.$$

Since $\mathcal{C}\left(U_t^i \mid s < t\right) = 0$ in the theoretical game, $\mathcal{RC}_t$ should be equal to 1 for any outcome $\mathbf{o}$. The smaller it is, the less important present and future actions are in determining the expected continuation payoff that agents may get.

Proposition 4 *(Asymptotic Indifference) Let* $\bar{u}_i = \max_{(a,x)\in A\times X} u_i\left[r_i\,(a, x_t)\right]$. *For all histories* h^∞ *and all* $a \in A$,

$$\mathcal{C}\left(\widehat{U}_{i,t}^{\mathbf{R}^{rcr}} \mid t\right)(\mathbf{o}) \leq \bar{u}^i\,(1-\beta)\,\eta\,(\beta, t) = \bar{u}^i \frac{1}{t+1} + o\left(\frac{1}{t}\right),$$

$$\mathcal{C}\left(\widehat{U}_{i,t}^{\mathbf{R}^{rcr}} \mid s > t\right)(\mathbf{o}) \leq \bar{u}^i\,(1-\beta)\,\beta\eta_\beta\,(\beta, t) = \bar{u}^i \frac{\beta^2}{t\,(1-\beta)} + o\left(\frac{1}{t}\right),$$

which implies that

$$\sup_{\mathbf{o},\mathbf{o}'\in(A\times X)^{\infty}}\left|\widehat{U}_{i,t}^{\mathbf{R}^{rcr}}\left(\mathbf{o}\mid h^{t}\right)-\widehat{U}_{i,t}^{\mathbf{R}^{rcr}}\left(\mathbf{o}'\mid h^{t}\right)\right|\leq o\left(\frac{1}{t}\right). \tag{8}$$

Moreover,

$$\mathcal{RC}_t\left(\mathbf{o}\right)\propto\frac{1}{1+\left(1-\beta\right)t}+o\left(1\right)$$

so all expressions converge to 0 as $t\to\infty$.

Proposition 4 illustrates how the contribution of present and future payoffs decreases in time until becoming negligible, so any decision taken at a late round will not greatly affect the expected utility the agent is going to get, as illustrated by condition (8). This is true both in absolute and relative terms: the contribution of present and future payoffs relative to already realized payoffs is almost inversely proportional to the round the agent is at.

4.1.2 Measuring distortions in implementability and payoffs

Equation (7) allows us to compare the incentive constraints quite easily. In each of the expressions, the present is evaluated in the same manner, $(1-\beta)\,u_i\,(r_{i,t})$, and the only differences come from discounting future payoffs, which is time dependent.

We now develop a measure of the distortion. Suppose we consider a constant outcome $(a_s,x_s)=(a,x)$ for all $s\geq t$ at time t, which generates a constant stream utility. The theoretical expected present value from any history h^t onwards, which we will denote by W_t, is

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

On the other hand, the when we do this computation in the game $\widehat{\Gamma}\left(\boldsymbol{R}_{rcr}\right)$, we have from (7) that,

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

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

$$\rho_t:=\frac{\widehat{W}_t}{W_t}=\left(1-\beta\right)\frac{\eta_\beta\left(\beta,t\right)}{\eta\left(\beta,t\right)}. \tag{9}$$

We show 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 (9) we are equipped with an explicit

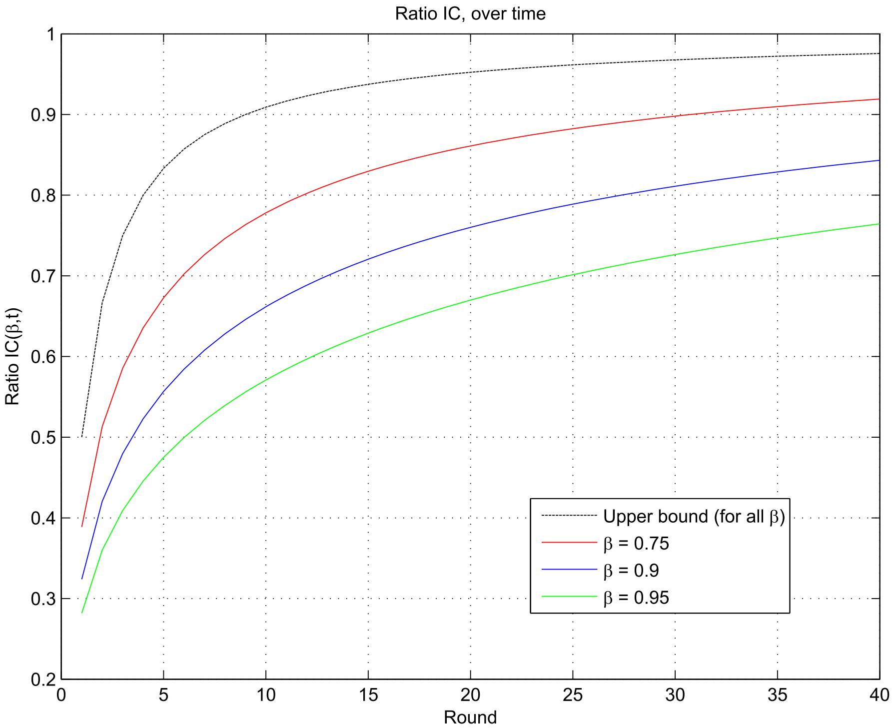

Fig. 1 Ratio of present values ρ_t versus round t

measure of how bad the problem is.[9] This is a measure of the distortion in the incentive condition.

Figure 1 explores the behavior of this ratio for different values of β, as t grows. For $\beta = 0.9$, at round 1 the ratio about 0.3, which implies that agents evaluate future relative utility streams at 30% the value in the target game, which gives a sizable measure of the distortion of incentives in (7). Even by round 10 this distortion is about 65%. In addition the figure displays a uniform bound across all discount factors.

To demonstrate how the slow convergence relates to the asymptotic indifference, we plot $\mathcal{RC}_t$ as well in Fig. 2. In the target game notice that the relative contribution is identically 1 at all periods. By studying the behavior of the ratio of present values ρ_t and the relative contribution together against time, we can see that as ρ_t slowly converges to 1 (and the distortion becomes arbitrarily smaller), meanwhile the relative contribution rapidly converges to zero. By period 10, the relative contribution has dropped to 0.44 and ρ_t is merely 0.67. This figure suggests that by the time that the valuation of relative future utility streams are close to the target game, agents are "almost indifferent" about the potential continuation histories they could face.

[9] We can calculate the functions η and η_β accurately using the finite integral formulation in Lemma 10.1 and 1.

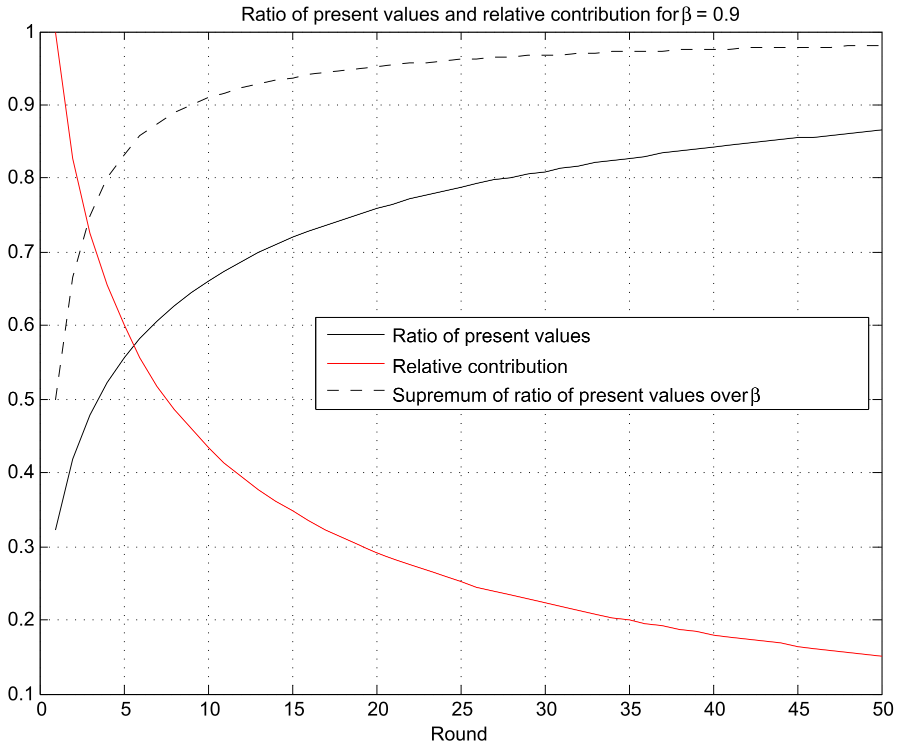

Fig. 2 ρ_t and $\mathcal{RC}_t$ versus round t

4.2 Payment for all rounds

In this Section we study the "all rounds: payment scheme and study under what conditions it may also fail to implement Γ given a utility function profile $\mathbf{u}$.[10] We establish that the payment schemes may significantly weaken the incentives of the participants as the number of rounds played increases and present a new result on uniform asymptotic indifference. It is well-known in the literature that the standard payment mechanism was generated with risk neutrality in mind (Murnighan and Roth 1983), though it is not uncommon to find examples that still use all round payment despite studying a model with curvature.

The payment scheme $\mathbf{R}^{all}$ is given by

$$R_{i,t}(h^t) = \sum_{s=0}^{s=t-1} r_i(a_s, x_s) \text{ with prob.1}$$

where $T \sim \text{Geom}(\beta)$. There may be a distortion due to $\mathbf{R}^{all}$. Since agents receive the payment for the experiment only when the experiment ends, the amount earned up to

[10] This concern has been raised by researchers interested in experiments studying behavior under risk-aversion: e.g., Charness and Genicot (2009) and Fischer (2013).

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 the long run. Note that if utility was linear in earnings, this payment scheme would not cause problems, as noted in the literature.

We formalize these intuitions in Proposition 5. Let

$$\overline{r}_i = \max_{(a,x)\in A\times X} r_i\,(a,x) \text{ and } \underline{r}_i = \min_{(a,x)\in A\times X} r_i\,(a,x)$$

be the best and worst possible stage rewards for agent i, and suppose that $\underline{r}_i \geq 0$.

Proposition 5 *Suppose that $u_i\,(\cdot)$ is an increasing, concave and differentiable function.*

1. *The range of values for contemporaneous and continuation utilities is decreasing over time. Specifically, for any feasible h^t and any pair of feasible continuation sequences $\mathbf{o}, \mathbf{o}' \in (A \times X)^{\infty}$ we have that*

$$\left|\widehat{U}_{i,t}^{\mathbf{R}^{all}}\left(\mathbf{o} \mid h^t\right) - \widehat{U}_{i,t}^{\mathbf{R}^{all}}\left(\mathbf{o}' \mid h^t\right)\right| \leq \frac{\beta}{1-\beta} u_i'(\underline{r}_i t)(\overline{r}_i - \underline{r}_i)$$

If, in addition, u_i satisfies the Inada condition $u_i'\,(\infty) = 0$ and $\underline{r}_i > 0$, then as $t \to \infty$

$$\sup_{\mathbf{o},\mathbf{o}'\in(A\times X)^{\infty}} \left|\widehat{U}_{i,t}^{\mathbf{R}^{all}}\left(\mathbf{o} \mid h^t\right) - \widehat{U}_{i,t}^{\mathbf{R}^{all}}\left(\mathbf{o}' \mid h^t\right)\right| \to 0. \quad (10)$$

2. *If u is linear (so agents are risk-neutral) then $\mathbf{R}^{all}$ implements $\Gamma_{\mathbf{u}}$*

Proposition 5 illustrates the nature of the distortion caused by $\mathbf{R}^{all}$. As time passes the amount by which an agent's utility changes must be decreasing; we note the parallel between (8) and (10). Note that the payment scheme is only implementing if the participants are modeled as risk-neutral, which of course is well-understood.

5 A model of borrowing and savings

In this section, we consider a simple environment: one agent who faces stochastic income and has access to savings and borrowing. This environment is a basic building block of numerous economic models and presents a simple one-person dynamic game.

Additionally, this example addresses the following issue. Observe that in a repeated game if agents can renegotiate in the future, since late rounds may have IC conditions that converge to the target IC conditions as $t \to \infty$, late round outcomes in the induced game may resemble those of the target game. This, however, is not at all robust: it will not hold in even the simplest dynamic game with a Markovian variable affected by past period interaction. This simplest setting is where one agent faces deterministic income and has a savings and borrowing vehicle.

The setup is entirely standard. An agent with preferences $U=(1-\beta)\sum_{t=0}^{\infty}\beta^{t}u(c_t)$ where $u'>0$ and $u''<0$, 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 gross interest rate $Q>0$. The budget constraint and no-Ponzi conditions are

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

Arrow's time-zero constraint is

$$\sum_{t=0}^{\infty}Q^{1-t}c_t=\sum_{t=0}^{\infty}Q^{1-t}y_t. \tag{11}$$

The usual Euler equation is $u'(c_t)=\beta Qu'(c_{t+1})$. In particular $\beta Q=1$ yields $c_t=c_{t+1}=c^*$. The consequence of this, of course, is Friedman's permanent income hypothesis which follows by (11), with $\sum_{t=1}^{\infty}\beta^{t-1}c^*=\sum_{t=1}^{\infty}\beta^{t-1}y_t\iff c^*=(1-\beta)\sum_{t=1}^{\infty}\beta^{t-1}y_t$. This is independent of the specific preferences we considered.

We show that even though in the limit both random round and all round payment generate the same inter-temporal tradeoff in late rounds, the actual consumption paths will be exceedingly different relative to the target game. Instead of exhibiting constant cross-period consumption, they will exhibit decaying consumption paths.

5.1 Payment for a randomly chosen round

Suppose that we pay the agent for a randomly chosen round in the above environment with $\beta Q=1$. It can be shown, using the recursive method shown in Appendix 1, that the modified Euler equation is

$$u'(c_t^{rar})=\underbrace{\frac{\eta(\beta,t+1)}{\eta(\beta,t)}}_{<1}\beta Qu'(c_{t+1}^{rar})<\beta Qu'(c_{t+1}^{rar})$$

which clearly distorts the natural Euler equation from the original model. Instead of consuming a constant amount, the agent would choose a forever decreasing consumption bundle. To further our intuition, define $\hat{Q}_t=\eta(\beta,t)^{-1}\cdot\eta(\beta,t+1)\,Q$ as the effective gross interest rate, as the Euler equation is $u'(c_t)=\beta\hat{Q}_t u(c_{t+1})$. Observe $\hat{Q}_t<Q$ for all periods; agents will save less under round at random payment than under the theoretical model.

Proposition 6 *Let $u(\cdot)$ be a strictly concave and differentiable utility function. Then*

$$\frac{u'(c_t^*)}{u'(c_t^{rar})}\propto\eta(\beta,t)\to 0 \text{ as } t\to\infty. \tag{12}$$

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

$$\frac{c_t^{rar}}{c_t^*} \propto [\eta(\beta, t)]^{\frac{1}{\sigma}} = O\left(t^{-\frac{1}{\sigma}}\right).$$

This implies that consumption under the randomly chosen round payment scheme is infinitely smaller than the theoretical predicted consumption decision when $t \to \infty$. Suppose now that $\beta Q = 1$, so $c_t^* = c^*$ for all t and suppose $\sigma = 1$. Then,

$$c_t = \frac{\eta(\beta, t)}{\eta(\beta, 1)} c_1 \propto \eta(\beta, t)$$

so $c_t \to 0$. Even if the Euler equation does converge to that 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 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. Figure 3 presents this result for various values of β. We have utilized log utility and set Q such that $Q\beta = 1$ in each case. Observe that instead of maintaining constant consumption over time, consumption steeply declines. This exhibits a substantive quantitative and qualitative differences in induced behavior.

5.2 Payments for all rounds

We consider the case where agents are paid for all rounds. In Appendix 1 we show that the modified Euler equation is now

$$u'(c_t^{all}) = \beta(Q-1)\, u'(c_t^{all}) < \beta Q u'(c_{t+1}^{all}) \tag{13}$$

Again this has the effect of reducing incentives for saving, since the effective net interest rate for the agent is $q := Q - 1$. The extent of the distortion is illustrated by the following proposition.

Proposition 7 *Let $u(\cdot)$ be a strictly concave and differentiable utility function. Then*

$$\frac{u'(c_t^*)}{u'(c_t^{all})} \propto \left(1 + \frac{1}{q}\right)^{-t}. \tag{14}$$

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

$$\frac{c_t^{all}}{c_t^*} \propto \left(\frac{q}{1+q}\right)^{\frac{t-1}{\sigma}} \to 0 \text{ as } t \to \infty$$

so no matter how small σ is, the ratio between consumptions goes to zero. In particular, consider the case where $1 < Q < 1 + \frac{1}{\beta}$. This implies that $\beta Q > 1$ and $\beta(Q-1) < 1$.

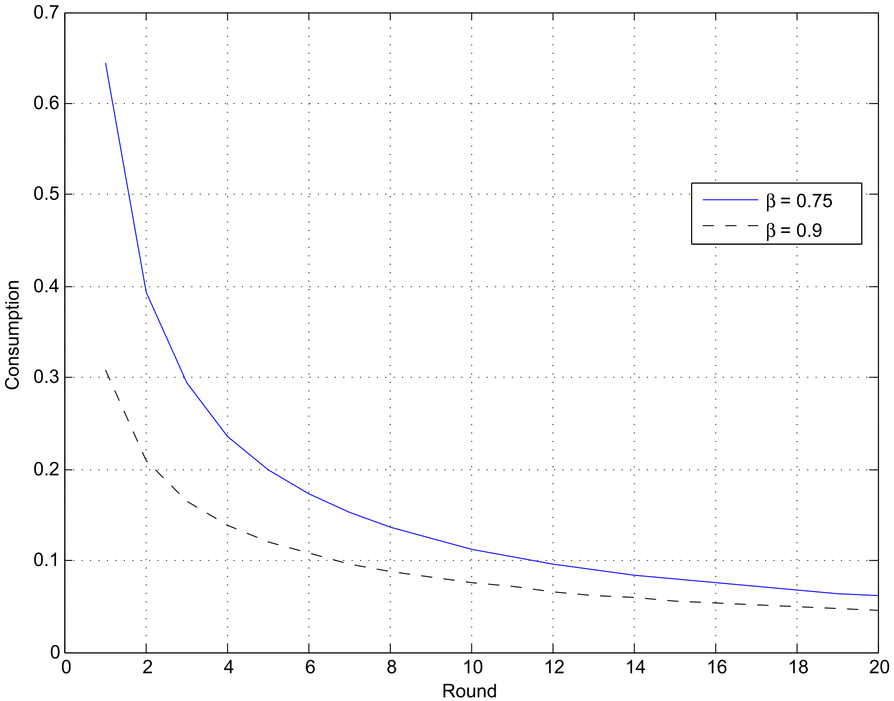

Fig. 3 Consumption sequences with log utility for payment for a randomly chosen round. Q set such that $\beta Q = 1$ and $y_t = 1$

The above equations imply that for any $\sigma > 0$ we have that

$$c_t^{all} \propto [\beta (Q-1)]^{\frac{t-1}{\sigma}} \to 0 \text{ since } \beta (Q-1) < 1$$

but

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

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, 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. Moreover, if the researcher is interested in larger-stakes games (see, e.g., Gneezy and Rustichini, 2000), it may very well be that curvature enters decisions thereby invalidating all round payment as an alternative to random round payment.

Figure 4 illustrates this, with $\beta = 0.97$, $Q = 1.3$ and log utility. Under the model, the agent should exhibit a growing consumption sequence. However, under all round

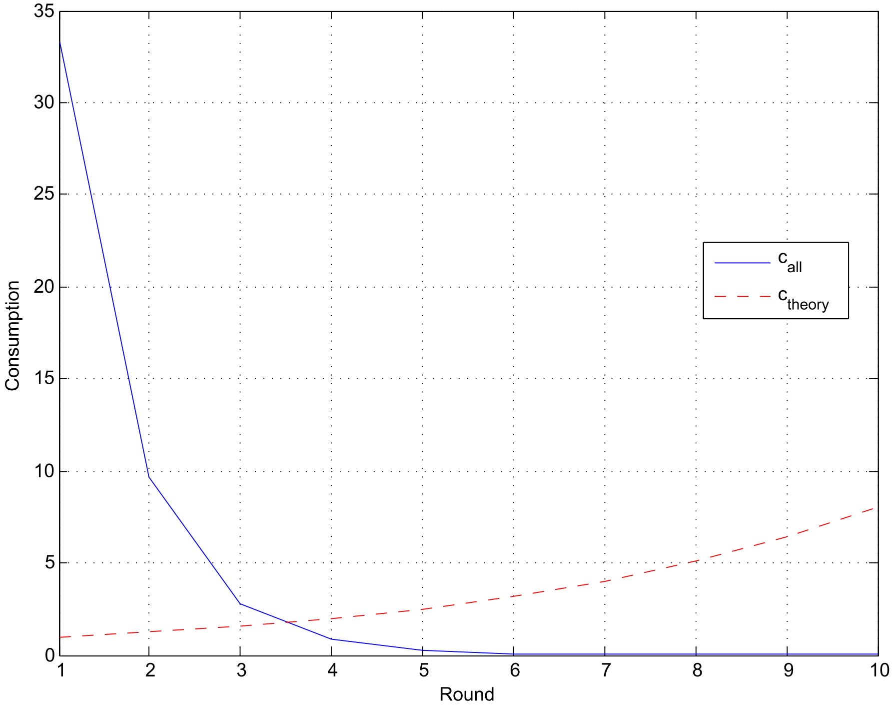

Fig. 4 Induced equilibrium consumption sequences under the theory and all round payment

payment, since all future streams of consumption until the end of the game is kept by the agent, the she wants to maximize the amount of consumption early on. Again we see large quantitative and qualitative disparities between the theoretical model and the induced game under incorrect payment.

6 Conclusion

We rigorously study payment schemes in multi-stage games and are interested in which schemes robustly implement the game in the sense that along all histories the preference orderings of all agents among all actions available coincide. This includes both settings where agents play or do not play equilibrium behavior. We begin by studying last round payments and show not only does it always implement the game of interest but, also, it is the only scheme to uniquely do so in a robust fashion. That is, last round payment is the only scheme that implements a given game for all preferences—which of course cannot be determined by the experimenter and is a fundamental feature of the human subject. We then study two schemes used in the literature: randomly chosen round and all round payment. We show that both schemes often will fail at implementing the target game that the experimentalist is interested in. The distortions can be quantitatively large and, moreover, can change the qualitative dynamics of behavior in a substantive way. In addition to generating distortions, these payment schemes are not robust in two ways. First, asymptotically, subjects should care less and less about their decisions, which makes the realized behavior not perturbation robust. Second, even arbitrarily small amounts of curvature can make all round payments fail. Last round payment never exhibits either of these issues. Ultimately, we argue that researchers should use last round payment schemes to implement infinite horizon dynamic games that satisfy our rather general class of assumptions.

Acknowledgements We thank Al Roth for extremely helpful discussions. Essential feedback was provided by Abhijit Banerjee, Ben Brooks, Gary Charness, Juan Dubra, Alex Frankel, Simon Gächter, Matt Jackson, Emir Kamenica, Anton Kolotilin, Jacob Leshno, Stephen Morris, Juan Passadore, Muriel Niederle, Charlie Sprenger, Jusso Toikka, Ivan Werning, and Xiao Yu Wang. Chandrasekhar thanks the NSF GRFP and the Alfred P. Sloan Foundation. This paper was previously circulated as "A Note on Payments in Experiments of Infinitely Repeated Games with Discounting".

Appendix A General analysis

A.1 Setup

We extend our results to multi-stage games. As we did in the dynamic game setting above, it is convenient to express such a game in two parts: a physical environment (which will be replicated in the experiment) and preferences over outcomes (that will be assumed in the theoretical model the economist chooses). We will do it for complete information games, although the extension to incomplete information games is straightforward.

A.1.1 Environment

A multi-stage environment E is described by a set of players I and an infinite number of stages $\tau \in \mathbb{N}$. For every history up to stage $h^{\tau-1}$, at stage τ agents play a finite extensive-form game

$$E\left(h^{\tau-1}\right) := \left\{\hat{\mathcal{H}}, \hat{\mathcal{Z}}, \left\{A_i\left(\hat{h}\right), I_i\left(\hat{h}\right)\right\}_{i\in[n],\hat{h}\in\mathcal{H}}, \left\{\hat{r}_i\left(z\right)\right\}_{z\in\hat{\mathcal{Z}}}\right\}$$

where:

- $\hat{\mathcal{H}} := \hat{\mathcal{H}} \mid h^{\tau-1}$ is the set of **partial stage histories**; i.e, partial histories of $E\left(h^{\tau}\right)$.
- $\hat{\mathcal{Z}} := \mathcal{Z} \mid h^{\tau-1}$ is the set of **partial stage terminal nodes**
- $A_i\left(\hat{h}\right) := A_i\left(\hat{h} \mid h^{\tau-1}\right)$ is a set of **stage actions** that agent i can make at history $\hat{h}$ of environment E, or likewise, the actions that agent i can take at history $\left(h^{\tau}, \hat{h}\right)$. When $A_i = \{a_i\}$ (singleton) then we say agent i is inactive at stage $\hat{h}$.
- $I_i\left(\hat{h}\right) := I_i\left(\hat{h} \mid h^{\tau-1}\right) \subseteq \hat{\mathcal{H}}$ is the **stage information set** of agent i at stage history $\hat{h}$.

- $\hat{r}_i(z) := \hat{r}_i\left(z \mid h^{\tau-1}\right)$ is the monetary reward function for stage τ, at terminal node z.

Also, denote by $\Sigma\left(h^{\tau-1}\right)$ the set of mixed contingent strategies for environment $E\left(h^{\tau-1}\right)$. We can incorporate any sort of random shocks simply by taking a player named "nature", who plays a known mixed strategy over her actions, which is measurable with respect to the observed public history. A history (unlike a stage history) is the description of the outcomes at all previous stages:

$$h = (z_1, z_2, \ldots, z_\tau)$$

where $z_k \in \mathcal{Z} \mid h^k$ for all $k \in [\tau]$. Let $l(h) = \tau$ be the length of history h. The set of all histories is written as $\mathcal{H}$ (unlike stage partial histories, written as $\hat{\mathcal{H}}$). For two given histories $h', h \in \mathcal{H}$ we write "h' is a successor of h" as $h' \succ h$. We also write Σ_i for the set of (mixed) contingent strategies $\sigma_i(h^\tau) \in \Sigma_i(h^\tau)$ for all $h^\tau \in \mathcal{H}$, and $\Sigma := \prod_{i\in[n]} \Sigma_i$.

We can then define rewards as functions of histories: if h' is the direct successor of history h; i.e, $h' = (h, z)$ for some $z \in \mathcal{Z} \mid h$, we define $r_i\left(h'\right) := \hat{r}_i(z \mid h)$

A.1.2 Preferences

Given the physical environment Γ, the experimenter needs to choose a preferences for agents, under which they rank which outcomes they prefer. A **preference model u** is determined by a set of **conditional payoff functions** $\mathbf{u} = \left\{U_i\left(h' \mid h\right)\right\}_{i\in[n],h,h'\in\mathcal{H}}$ where $h' \succ h$. We say that a preference model $\mathbf{U}$ has **common discounting** if and only if there exist functions $u_i : \mathbb{R} \to \mathbb{R}$ and a process $\mathbf{b} = \{b(h)\}_{h\in\mathcal{H}} \geq 0$ such that $\sum_{h\in\mathcal{H}} b(h) = 1$ for all $h \in \mathcal{H}$ and for all $h' \succeq h$:

$$U_i\left(h' \mid h\right) = \sum_{\hat{h}:h\preceq\hat{h}\preceq h'} b\left(\hat{h}\right) u_i\left[r_i\left(z_{l\left(\hat{h}\right)} \mid \hat{h}\right)\right]. \tag{A1}$$

In particular, we say that a model $\mathbf{U}$ has common exponential discounting when $\exists\beta \in (0, 1)$ such that

$$U_i\left(h' \mid h\right) = (1-\beta) \sum_{h'':h'\succeq h''\succeq h} \beta^{l\left(\hat{h}\right)} u_i\left[r_i\left(z_{l\left(\hat{h}\right)} \mid \hat{h}\right)\right]. \tag{A2}$$

where $l(h) = \{$ length of history$\}$. From now on, we will only consider common discounting models: let $\mathcal{U}(\Gamma, \mathbf{b})$ be the set of common exponential discounting models given an environment Γ.

Finally, given a profile of monetary utility functions $\mathbf{u} = (u_i)_{i\in I}$ we define the **multi-stage game** $\Gamma_{\mathbf{u}}$ of environment E and preferences over outcomes for all agents given by (A2).

A.2 Implementation

Given an exponential discounting process (parametrized by $\beta \in (0, 1)$) and given a profile of utility functions $\mathbf{u}$, define the associated game $\hat{\Gamma}_{\mathbf{u}}(\mathbf{R})$ as the associated finite game that satisfies the following:

- Game $\Gamma\left(h^0\right)$ is played,
- With probability β the game continues and agents play game $\Gamma_{\mathbf{u}}\left(h^1\right)$, and with probability $1-\beta$ the game ends.
- After k stages, game ends with probability $1-\beta$ or the game continues for at least one more stage, playing now the extensive form game $\Gamma_{\mathbf{u}}\left(h^{k-1}\right)$.
- When game ends, agents receive a random payment, according to distribution $\mathbf{R}\left(h^T\right)$ (where T is the last stage game) and agents then get $\hat{U}_i\left(h^T\right) = E_T\left\{E_r\left[u_i\left(r_i\right) \mid h^T\right]\right\}$ where the expectation $E_r(\cdot \mid h^T)$ is calculated using the probability measure $\mathbf{R}\left(h^T\right)$

In particular, we generalize the last round payment as $\mathbf{R}_i^{last}\left(h^T\right) = r_i\left(z_T, h^{T-1}\right)$ with probability 1. For general discounting processes $b\left(h\right)$ we modify the above method, by changing the probability of the game ending in each particular history. The associated game $\hat{\Gamma}_{\mathbf{u}}(\mathbf{R})$ is played as follows:

- Game $\Gamma\left(h^0\right)$ is played.
- With probability $1-\hat{\beta} = b\left(h^0\right)$ the game ends after the first round, and with probability $\hat{\beta} = 1 - b\left(h^0\right)$ the game continues for another round.
- At history $h = h^t$ the game ends with probability

$$1 - \hat{\beta} := \frac{b\left(h\right)}{1 - \sum_{h' \prec h} b\left(h'\right)}$$

  i.e, given the fact that in all previous periods the game did not end.

See that this process implies that Pr (game ends at h) $= b\left(h\right)$, in which case the agent would obtain $u_i\left(r_i\left(h\right)\right)$ utils, which then trivially extends the result of exponential discounting models.

Proposition 8 *(Strong Implementation) The payment scheme* $\mathbf{R}^{last}$ ***strongly implements*** *environment* E.

A.3 Payoff invariant solutions

We now discuss a different approach, in which the researcher only wants to check whether a particular solution to the game will be played, which is expressed as a restriction over the set of feasible strategies that agents may play. We formalize this below

Definition 4 (Solution) A **solution** is a correspondence $\mathcal{S}\left(\Gamma\right) \subseteq \Sigma$ for all Γ based on E

Natural solution concepts may be the set of Subgame Perfect Equilibria, Nash Equilibria, Sequential Equilibria, and so on. We will focus on a particular class of solution concepts: *payoff invariant solutions*. These are solution concepts that are basically defined by the agents preferences alone, and where particular labels to strategies are not relevant. This class that encompass most solution concepts used in the literature.

We say that games $\Gamma_{\mathbf{u}}$, and $\Gamma_{\overline{\mathbf{u}}}$, are **payoff-equivalent** if for all $h \in \mathcal{H}$ and all $h', h'' \succ h \in \mathcal{H}$ we have

$$U_i\left(h' \mid h\right) \geq U_i\left(h'' \mid h\right) \iff \overline{U}_i\left(h' \mid h\right) \geq \overline{U}_i\left(h'' \mid h\right) \iff \text{ for all } i \in [n]. \tag{A3}$$

That is, both games have same histories, strategy spaces, and same preferences. With some abuse of notation, we write $\Gamma_{\mathbf{u}} \sim \Gamma_{\overline{\mathbf{u}}}$ for payoff-equivalent games.

Definition 5 (Payoff invariance) A solution $\mathcal{S}(\Gamma)$ is **payoff-invariant** if $\Gamma_{\mathbf{u}} \sim \Gamma_{\overline{\mathbf{u}}}$ implies $\mathcal{S}(\Gamma_{\mathbf{u}}) = \mathcal{S}(\Gamma_{\overline{\mathbf{u}}})$.

Most of the solution concepts we use are payoff-invariant. For example, the set of subgame perfect equilibria of a game is payoff invariant, since its definition depends solely on the conditional payoffs: let $h'(\sigma) \mid h \in \mathcal{H}$ be the continuation history according to strategy profile σ, conditional on reaching history h.

$$\mathbf{SPE}(\Gamma_{\mathbf{u}}) = \left\{\sigma \in \Sigma : U_i\left[h'(\sigma_i) \mid h\right] \geq U_i\left[h'\left(\hat{\sigma}_i, \sigma_{-i}\right) \mid h\right] \text{ for all } h \in \mathcal{H}, \hat{\sigma}_i \in \Sigma_i\right\}$$

Other examples of such solutions are include Nash equilibrium, Markov equilibria, best and/or worst NE, the core, Nash bargaining, and so on.

Proposition 9 *Let $\mathcal{S}$ be a payoff invariant solution. Then $\mathcal{S}(\Gamma_{\mathbf{u}}) = \mathcal{S}\left(\hat{\Gamma}_{\mathbf{u}}\left(\mathbf{R}^{last}\right)\right)$ for all* $\mathbf{u}$

Appendix B Proofs

Proof of Proposition 2 For any feasible outcome $\mathbf{o} = (a_s, x_s)_{s=1}^{\infty}$ define the outcomes $\tilde{\mathbf{o}} = (h^t\mathbf{o})$ and $\hat{\mathbf{o}} = (\hat{h}^t\mathbf{o})$.Take the histories $h^t, \hat{h}^t$ and the agent i that are implied by the fact that $\mathbf{R}$ disagrees with $\mathbf{R}^{last}$, so that $h^t \neq \hat{h}^t$ but $h_{t-1} = (a_{t-1}, x_{t-1}) = \hat{h}_{t-1}$. Let $R = R_i\left(h^t\right)$ be the distribution of payments according to history h^t, and $\hat{R} = R_i\left(\hat{h}^t\right)$ the corresponding distribution for history $\hat{h}^t$. Let $S = \mathbf{Supp}(R) = \{r_1, r_2, \ldots, r_m\}$ and $\hat{S} = \mathbf{Supp}\left(\hat{R}\right) = \left\{\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_n\right\}$ and $p \in \Delta(S)$ and $\hat{p} \in \Delta\left(\hat{S}\right)$ the corresponding probabilities for each monetary reward. Suppose first that $S \neq \hat{S}$, so that $\exists r_j \in S \cap \hat{S}^c$. Given a utility function u_i, we can then define the expected utilities $V_i^{\mathbf{R}}\left(h^t\right) := E_r\left[u_i(r) \mid h^t\right] = \sum_{r \in S} p(r)\, u_i(r)$ and the equivalent for $\hat{h}^t$.

Define $r_1^* = \max\left\{S \cup \hat{S}\right\}$ to be the maximum return. We will first construct utility functions of the form

$$\hat{u}(c) = f_k(c) = \begin{cases} 0 & \text{if } c < k \\ 1 & \text{if } c \geq k \end{cases}$$

to show that $V_i^{\mathbf{R}}(h^t) \neq V_i^{\mathbf{R}}(\hat{h}^t)$.

Case 1: $r_1^* \notin S \cap \hat{S}$: this is the case where the maximum reward is only achieved (with positive probability) only in one of the histories. Then, if we define $\hat{u}_i(c) = f^{r_1^*}(c)$ we will have that $V_i^{\mathbf{R}}(h^t) > V_i^{\mathbf{R}}(\hat{h}^t)$ when $r_1^* \in S$ or $V_i^{\mathbf{R}}(h^t) < V_i^{\mathbf{R}}(\hat{h}^t)$ if $r_1^* \in \hat{S}$.

Case 2: $r_1^* \in S \cap \hat{S}$. Here we need to consider several sub-cases: **(2.a)** $p(r_1^*) \neq \hat{p}(r_1^*)$ and **(2.b)** $p(r_1^*) = \hat{p}(r_1^*)$. For (**2.a**), the function $\hat{u}_i(c) = f^{r_1^*}(c)$ will still make $V_i^{\mathbf{R}}(h^t) > V_i^{\mathbf{R}}(\hat{h}^t)$ if and only if $p(r_1^*) > \hat{p}(r_1^*)$.

If we are in Case (**2.b**), then we need to define $r_2^* = \max\left\{r \in S \cup \hat{S} : r \neq r_1^*\right\}$ as the second best reward. Then, we replicate the analysis of Case 1 for $r = r_2^*$ and Case (**2.a**). If we get that $r_2^* \in S \cap \hat{S}$ and $p(r_2^*) = \hat{p}(r_2^*)$ we then go to the third best r_3^* and proceed iteratively. Eventually, we either stop at $r = r_m^*$ for $m < \#(S \cup \hat{S})$ or run out of rewards in both supports. However, this only happens if $S = \hat{S}$ and $p(r) = \hat{p}(r) \forall r$, which would violate the assumption that $\mathbf{R}(h^t) \neq \mathbf{R}(\hat{h}^t)$.

Once we find r^* we also find a candidate utility function $\hat{u}_i(c) = f_{r^*}(c)$, which is non-decreasing (but not strictly increasing) and satisfies $\sum_{s \in S} u_i(r) p(r) > \sum_{s \in \hat{S}} u_i(r) q(r)$ (or vice versa) given that this inequality holds if and only if $\sum_{r \geq k^*} p(r) \geq \sum_{r \geq k^*} q(r)$. Consider the function $g_{r,\eta}(c) = \Phi\left((c - r + \epsilon)/\eta\right)$ where $\Phi(\cdot)$ is the cdf of the normal distribution $\mathcal{N}(0, 1)$ and $\eta > 0$. It is easy to see that, for all $c \neq r$ we have that $\lim_{\eta \to 0} g_{r,\eta}(c) = 1$ when $c > r$ and $\lim_{\eta \to 0} g_{r,\eta}(c) = 0$ when $c < r$, so $g_{r,\eta}(c)$ converges point-wise to $f_r(c)$. Let $\epsilon = \alpha \min_{r \in S \cup \hat{S}} |r^* - r|$ with $\alpha < 1$ and define $r = r^* - \epsilon$ (so there is no reward in the support of either distribution such that $c = r^*$). We then have that $\sum_{r \in S} g_\eta(r) p(r) \to \sum_{r \in S : r \geq r^*} p(r) = V_i^{\mathbf{R}}(h^t)$ as $\eta \to 0$, and the same holds for $V_i^{\mathbf{R}}(\hat{h}^t)$. Therefore, by choosing a small enough η, we construct the utility function $u_i(c) := g_{r^* - \epsilon, \hat{\eta}}(c)$ that satisfies that either $V_i^{\mathbf{R}}(h^t) > V_i^{\mathbf{R}}(\hat{h}^t)$ or $V_i^{\mathbf{R}}(h^t) < V_i^{\mathbf{R}}(\hat{h}^t)$, showing the first result.

For the second result, we take the case where $V_i^{\mathbf{R}}(h^t) > V_i^{\mathbf{R}}(\hat{h}^t)$. Since $g_{r,\eta}(c)$ is strictly increasing and $\mathbf{R}$ satisfies quasi-separability, we then have that $V_i^{\mathbf{R}}(h^t h^s) > V_i^{\mathbf{R}}(\hat{h}^t h^s)$ for any partial continuation history h^s. For any outcome $\mathbf{o} = (\hat{a}_s, \hat{x}_s)_{s=0}^{\infty}$ where $(\hat{a}_0, \hat{x}_0) = (a_{t-1}, x_{t-1})$ (i.e, it is consistent with both h^t and $\hat{h}^t$) define $h_{\mathbf{o}}^s := (a_k, x_k)_{k=0}^{k=s-1}$ as the partial history of its first s rounds. Then, for any such outcome, we have that:

$$\hat{U}_i^{\mathbf{R}}(\mathbf{o} \mid h^t) - \hat{U}_i^{\mathbf{R}}(\mathbf{o} \mid \hat{h}^t) = (1 - \beta)\left[V_i^{\mathbf{R}}(h^t) - V_i^{\mathbf{R}}(\hat{h}^t)\right]$$
$$+\beta(1 - \beta)\sum_{s=0}^{\infty} \beta^s \left[V_i^{\mathbf{R}}(h^t h_{\mathbf{o}}^s) - V^{\mathbf{R}}(\hat{h}^t h_{\mathbf{o}}^s)\right] > 0$$

using the fact that $V_i^{\mathbf{R}}(h^t h_{\mathbf{o}}^s) > V_i^{\mathbf{R}}(\hat{h}^t h_{\mathbf{o}}^s)$ for all $s \geq 0$, implying then that $\hat{U}_i^{\mathbf{R}}(\mathbf{o} \mid h^t) > \hat{U}_i^{\mathbf{R}}(\mathbf{o} \mid \hat{h}^t)$, as we wanted to show. □

Proof of Corollary 3 If $\mathbf{R}$ disagrees with $\mathbf{R}^{last}$ at $h^t, \hat{h}^t$, we can choose t to be the first round when $\mathbf{R}$ disagrees with $\mathbf{R}^{last}$ (so, before that, they have the same expected values). Let for $s \leq t$ let h^s and $\hat{h}^s$ be the partial histories up to s of each history. Take any outcome $\tilde{\mathbf{o}}$ consistent with (a_{t-1}, x_{t-1}) and define the outcomes $\mathbf{o} = (h^t \tilde{\mathbf{o}})$ and $\hat{\mathbf{o}} = (\hat{h}^t \tilde{\mathbf{o}})$. From the time $t = 0$ perspective, we have that

$$U_i(\mathbf{o} \mid x_0) = (1-\beta) \sum_{s=0}^{s=t-1} \beta^s u_i\left[r_i(a_s, x_s)\right] + (1-\beta)\beta^t \sum_{k=0}^{\infty} \beta^k u_i(\tilde{a}_k, \tilde{x}_k)$$
$$U_i(\hat{\mathbf{o}} \mid x_0) = (1-\beta) \sum_{s=0}^{s=t-1} \beta^s u_i\left[r_i(\hat{a}_s, \hat{x}_s)\right] + (1-\beta)\beta^t \sum_{k=0}^{\infty} \beta^k u_i(\tilde{a}_k, \tilde{x}_k)$$

and therefore

$$U_i(\mathbf{o} \mid x_0) - U_i(\hat{\mathbf{o}} \mid x_0) = (1-\beta) \sum_{s=0}^{s=t-1} \beta^s \left\{u_i\left[r_i(a_s, x_s)\right] - u_i\left[r_i(\hat{a}_s, \hat{x}_s)\right]\right\}$$

i.e. only the first t periods should matter. However, under $\mathbf{R}$ we have that for all $s \leq t$ we have $E^{\mathbf{R}}(u_i(r) \mid h^s) = u_i\left[r_i(a_s, x_s)\right]$ and $E^{\mathbf{R}}(u_i(r) \mid \hat{h}^s) = u_i\left[r_i(\hat{a}_s, \hat{x}_s)\right]$. Therefore

$$\begin{aligned} &U_i^{\mathbf{R}}(\mathbf{o} \mid x_0) - U_i^{\mathbf{R}}(\hat{\mathbf{o}} \mid x_0) = U_i(\mathbf{o} \mid x_0) - U_i(\hat{\mathbf{o}} \mid x_0) \\ &\quad + (1-\beta)\beta^t \left[\sum_{k=0}^{\infty} \beta^k \left(U_i^{\mathbf{R}}(\tilde{\mathbf{o}} \mid h^t h^s) - U_i^{\mathbf{R}}(\tilde{\mathbf{o}} \mid \hat{h}^t h^s)\right)\right] > U_i(\mathbf{o} \mid x_0) - U_i(\hat{\mathbf{o}} \mid x_0) \end{aligned}$$

This implies that if $U_i(\mathbf{o} \mid x_0) \geq U_i(\hat{\mathbf{o}} \mid x_0)$ we must also have $U_i^{\mathbf{R}}(\mathbf{o} \mid x_0) \geq U_i^{\mathbf{R}}(\hat{\mathbf{o}} \mid x_0)$, but it is not implied in the other direction. See that this result holds for any continuation outcome $\tilde{\mathbf{o}}$, so the difference $U_i^{\mathbf{R}}(\mathbf{o} \mid x_0) - U_i^{\mathbf{R}}(\hat{\mathbf{o}} \mid x_0)$ depends crucially on $\tilde{\mathbf{o}}$ while the theoretical comparison does not. □

Proof of Proposition 5 Observe that

$$\begin{aligned} \widehat{U}_i^{\mathbf{R}_{all}}\left(\mathbf{a} \mid h^t\right) - \widehat{U}_i^{\mathbf{R}_{all}}\left(\mathbf{a}' \mid h^t\right) = (1-\beta) \sum_{s=0}^{\infty} \beta^s \Bigg[& u_i\left(R_{i,t-1} + \sum_{k=0}^{k=s} r_{t+k}\right) \\ & - u_i\left(R_{i,t-1} + \sum_{k=0}^{k=s} r'_{t+k}\right)\Bigg] \end{aligned}$$

where $r_{i,t+k} = r_i\,(a_{t+k})$ 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

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

Using $x = R_{i,t-1} + \sum_{k=0}^{s} r_{t+k}$ and $y = R_{i,t-1} + \sum_{k=0}^{s} r'_{t+k}$,

$$(1-\beta)\sum_{s=0}^{\infty}\beta^s u'_i\left(R_{i,t-1}+\sum_{k=0}^{s} r_{i,t+k}\right)\times\left(\sum_{k=0}^{s}\left(r_{i,t+k}-r'_{i,t+k}\right)\right)\leq \widehat{U}_i^{\mathbf{R}_{all}}\left(\mathbf{a}\mid h^t\right)-\widehat{U}_i^{\mathbf{R}_{all}}\left(\mathbf{a}'\mid h^t\right)$$
$$\leq (1-\beta)\sum_{s=0}^{\infty}\beta^s u'_i\left(R_{i,t-1}+\sum_{k=0}^{s} r'_{i,t+k}\right)\times\left(\sum_{k=0}^{s}\left(r_{i,t+k}-r'_{i,t+k}\right)\right)$$

As u_i is concave, $u'_i\left(R_{i,t-1} + \sum_{k=0}^{s} r_{i,t+k}\right) \leq u'\left(R_{i,t-1}\right)$ and the same for $r'_{i,t+k}$, so

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

If $\underline{r}_i > 0$, then $R_{i,t} \geq \underline{r}_i(t+1)$, implying that

$$\left|\widehat{U}_i^{\mathbf{R}_{all}}\left(\mathbf{a}\mid h^t\right)-\widehat{U}_i^{\mathbf{R}_{all}}\left(\mathbf{a}'\mid h^t\right)\right| < u'_i(\underline{r}_i(t+1))\left[(1-\beta)\sum_{s=0}^{\infty} s\beta^s\left|\overline{r}_i-\underline{r}_i\right|\right]$$
$$\leq \frac{\beta}{1-\beta}u'_i(\underline{r}_i(t+1))\left(\overline{r}_i-\underline{r}_i\right)$$

as we wanted to show, using the fact that $\sum_{s=0}^{\infty} s\beta^s = \beta/(1-\beta)^2$. To show the contemporary utility result, without loss of generality take $r, \widehat{r} \in r_i(A)$, such that $r > \widehat{r}$. Then,

$$u'_i\left(R_{i,t-1}+\overline{r}_i\right)(r-\widehat{r}) \leq u'_i\left(R_{i,t-1}+r\right)(r-\widehat{r}) \leq (1-\beta)\,u_i\left(R_{i,t-1}+r\right)$$
$$-(1-\beta)\,u_i\left(R_{i,t-1}+\widehat{r}\right)$$
$$\leq u'_i\left(R_{i,t-1}+\widehat{r}\right)(r-\widehat{r}) \leq u'_i\left(R_{i,t-1}+\overline{r}_i\right)\left(\overline{r}_i-\underline{r}_i\right)$$

which implies that

$$\left|u_i\left(R_{i,t-1}+r\right)-u_i\left(R_{i,t-1}+\hat{r}\right)\right| \leq u'_i\left(R_{i,t-1}+\underline{r}_i\right)\left(\overline{r}_i-\underline{r}_i\right) \leq u'_i\left(\underline{r}_i t\right)\left(\overline{r}_i-\underline{r}_i\right).$$

as we wanted to show. □

Proof of Proposition 4 From Lemma 12, equation C7, we have that for any (a_t, x_t) the contribution of t period strategies is $\mathcal{C}\left(U_t^i \mid t\right)(a_t, x_t) = (1-\beta)\,\eta\,(\beta, t)\,u_i\,[r_i\,(a_t, x_t)]$ and the result follows from there. Also, Property (1) in Lemma 10 gives the approximation result (and convergence to zero). For any history a_{t+s}, using again equation (C7), we have

$$\mathcal{C}\left(U_t^i \mid \{s \geq t+1\}\right)\left(\{a_{t+s}, x_{t+s}\}_{s=1}^{\infty}\right) = \beta\,(1-\beta)\sum_{s=0}^{\infty}\eta\,(\beta, t+1+s)\,u_i\left[r_i\,(a_{t+1+s}, x_{t+1+s})\right]$$

$$\leq \beta\bar{u}_i\,(1-\beta)\sum_{s=0}^{\infty}\sum_{k=s}^{\infty}\frac{\beta^k}{t+1+k} = \beta\bar{u}_i\,(1-\beta)\sum_{k=0}^{\infty}\sum_{s=0}^{k}\frac{\beta^k}{t+1+k}$$

$$= \beta\bar{u}_i\,(1-\beta)\sum_{k=0}^{\infty}\frac{(k+1)\,\beta^k}{t+(k+1)} = \beta\bar{u}_i\,(1-\beta)\left(\sum_{j=1}^{\infty}\frac{j\beta^{j-1}}{t+j}\right)$$

$$= \beta\bar{u}_i\,(1-\beta)\left(\sum_{j=0}^{\infty}\frac{j\beta^{j-1}}{t+j} - 0\right) = \beta\bar{u}_i\,(1-\beta)\,\eta_\beta\,(\beta, t)$$

and the approximation result comes from property (1) in Lemma 10. □

Proof of Proposition 6 Using the Euler equation for both the theoretical and the RCR sequence of optimal consumption, we get

$$u'\left(c_t^*\right) = \frac{1}{\beta^{t-1}Q^{t-1}}u'\left(c_1^*\right)$$

and

$$u'(c_t^{rcr}) = \frac{1}{\beta^{t-1}Q^{t-1}}u'(c_1^{rcr})\frac{\eta\,(\beta, 1)}{\eta\,(\beta, t)}$$

then

$$\frac{u'(c_t^*)}{u'(c_t^{rcr})} = \frac{u'(c_1^*)}{u'(c_1^{rcr})}\frac{\eta\,(\beta, t)}{\eta\,(\beta, 1)} \propto \eta\,(\beta, t)\,.$$

□

Proof of Proposition 7 From before we knew that

$$u'\left(c_t^*\right) = \frac{1}{\beta^{t-1}Q^{t-1}}u'\left(c_1^*\right)$$

and likewise, we show that

$$u'(c_t^{all}) = \frac{1}{\beta^{t-1}(Q-1)^{t-1}}u'(c_1^{all})$$

Therefore

$$\frac{u'(c_t^*)}{u'(c_t^{all})} = \left(\frac{Q-1}{Q}\right)^{t-1}\frac{u'(c_1^*)}{u'(c_1^{all})} \propto \left(1+\frac{1}{q}\right)^{-t}.$$

□

Appendix C Auxiliary results

C.1 Technical results

Lemma 10 *Let η and ϕ be defined as in 3 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 $\eta(\beta, t) = \frac{1}{\beta^t}\int_0^\beta \frac{z^t}{z(1-z)} dz$.
4. For all β, t we can write η_β as $\eta_\beta(\beta, t) = \sum_{k=0}^{\infty} \frac{k\beta^{k-1}}{t+k} = \frac{1}{\beta^{t+1}}\int_0^\beta \frac{z^t}{(1-z)^2} dz$.
5. For given β: $\eta(\beta, t) = \frac{1}{t(1-\beta)} + o\left(\frac{1}{t}\right)$ 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 (1). 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

$$\eta(\beta, t+1) = \frac{1}{\beta}\left[\eta(\beta, t) - \frac{1}{t}\right] \tag{C4}$$

and is easy to see that η as defined in 3 satisfies this recursion. To show that this recursion is true, we do some algebra:

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

For 1, see that

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

which is valid since η is a power series. For (1) we can also use this to differentiate η:

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

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 (1), 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 $f_t \nearrow \beta^k$ point-wise. 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 (1) 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 $\frac{tk}{t+k}\beta^k \nearrow k\beta^k$ point-wise, 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 (1) 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 11 *Define the function* $B(\beta,t,s)$*, as*

$$B(\beta,t,s) = \beta^s\frac{\eta(\beta,t+1+s)}{\eta(\beta,t)}. \tag{C5}$$

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*

$$B(\beta,t,s) \nearrow \beta^s \text{ for all } s, \text{ as } t \to \infty. \tag{C6}$$

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

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

$$= \frac{\left(\frac{\beta^s}{1-\beta}\right)}{\left(\frac{1}{1-\beta}\right)} = \beta^s.$$

In (i) we used 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 12 *After history h^t, the utility of agent i of stream of rewards $r_{i,t} := r_i(a_t, x_t)$ from outcome $\mathbf{o} = \{a_t, x_t\}_{t=0}^{\infty}$ in $\hat{\Gamma}(\boldsymbol{R}^{rcr})$ can be written as*

$$\hat{U}_i\left(\mathbf{o} \mid h^t\right) = (1-\beta)\,\eta(\beta, t) \sum_{k=0}^{t} u_i(r_{i,t}) + \beta \left\{(1-\beta) \sum_{s=0}^{\infty} \beta^s \eta(\beta, t+1+s)\, u_i(r_{i,t+s+1})\right\} \tag{C7}$$

Proof We can decompose the utility as

$$U_t^i = (1-\beta) \frac{1}{t+1} \sum_{k=0}^{k=t} u_i\left(r_{i,k}\right) + \beta W^i, \tag{C8}$$

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

$$\begin{aligned}
W^i &= (1-\beta)\sum_{j=0}^{\infty}\beta^j\frac{1}{t+1+j}\left[\sum_{k=0}^{k=t}u_i\left(r_{i,k}\right)+\sum_{k=0}^{j}u_i\left(r_{i,t+k+1}\right)\right]\\
&= (1-\beta)\sum_{j=0}^{\infty}\beta^j\frac{1}{t+1+j}\sum_{k=0}^{k=t}u_i\left(r_{i,k}\right)+(1-\beta)\sum_{j=0}^{\infty}\beta^j\frac{1}{t+1+j}\sum_{k=0}^{j}u_i\left(r_{i,t+k+1}\right)\\
&= (1-\beta)\,\eta\left(\beta,t+1\right)\sum_{k=0}^{k=t}u_i\left(r_{i,k}\right)+(1-\beta)\sum_{j=1}^{\infty}\sum_{k=0}^{j}\frac{\beta^j}{t+1+j}\\
&= (1-\beta)\,\eta\left(\beta,t+1\right)\sum_{k=0}^{k=t}u_i\left(r_{i,k}\right)+(1-\beta)\sum_{k=0}^{\infty}u_i\left(r_{i,t+k+1}\right)\left(\sum_{j=k}^{\infty}\frac{\beta^j}{t+1+j}\right)\\
&\iff W^i=(1-\beta)\,\eta\left(\beta,t+1\right)\sum_{k=0}^{k=t}u_i\left(r_{i,k}\right)\\
&\qquad+(1-\beta)\sum_{s=0}^{\infty}\beta^s\eta\left(\beta,t+s+1\right)u_i\left(r_{i,t+k+1}\right).
\end{aligned}\tag{C9}$$

Therefore, putting together equations (C8) and (C9) we get

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

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

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

as we wanted to show. □

C.2 Recursive method for RCR payment

The typical dynamic programming program involves solving

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

such that

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

The usual Bellman equation is

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

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

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

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

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

such that

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

We will define $H(x_t, x_{t+1}, \eta_t) = \eta_t F(x_t, x_{t+1})$. In addition, given β, $\eta(\beta, t)$ is a strictly decreasing function of t. Therefore, let $T(\eta, \beta)$ be the inverse.[11] Using (C4) and augmenting the state space with η, which has a known law of motion, yields

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

such that

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

The Bellman equation for this problem is simply

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

11 $T(\eta, \beta)$ satisfies $\eta(\beta, T(\eta, \beta)) = \eta$.

such that

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

C.3 Recursive method for all round payment

We want to get the Euler equation in (13). We can characterize the optimal allocation by means of the following Bellman equation:

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

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

$$\begin{aligned}(1-\beta) u'\left(S + y + Ra - a'\right) = \beta \{ & V_a\left(a', S + y + Ra - a', y'\right) \\ & - V_S\left(a', S + y + Ra - a', y'\right)\} \end{aligned} \tag{C10}$$

and by the envelope conditions we have

$$V_a(a, S, y) = R(1-\beta) u'\left(S + y + Ra - a'\right) \tag{C11}$$

$$V_S(a, S, y) = (1-\beta) u'\left(S + y + Ra - a'\right). \tag{C12}$$

Using (C10) and substituting for c_t and c_{t+1}, we then get that

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

as we wanted to show.

References

Azrieli, Y., Chambers, C.P., Healy, P.J.: Eliciting multiple choices: A theoretical analysis of incentives in experiments. Technical report, Working Paper, Ohio State University (2012)

Azrieli, Y., Chambers, C.P., Healy, P.J.: Incentives in experiments: a theoretical analysis. J. Polit. Econ. **126**(4), 1472–1503 (2018)

Brown, A.L., Chua, Z.E., Camerer, C.F.: Learning and visceral temptation in dynamic saving experiments. Q. J. Econ. **124**(1), 197–231 (2009)

Cabral, L., Ozbay, E.Y., Schotter, A.: Intrinsic and instrumental reciprocity: An experimental study. Technical report, Technical report, Working paper. New York University (2011)

Chandrasekhar, A.G., Kinnan, C., Larreguy, H.: Informal insurance, social ties, and financial development: Evidence from a lab experiment in the field. In: MIT Working Paper (2011)

Charness, G., Genicot, G.: Informal risk sharing in an infinite-horizon experiment*. Econ. J. **119**(537), 796–825 (2009)

Cooper, R., DeJong, D.V., Forsythe, R., Ross, T.W.: Forward induction in the battle-of-the-sexes games. Am. Econ. Rev. **5**, 1303–1316 (1993)

Dal Bo, P.: Cooperation under the shadow of the future: experimental evidence from infinitely repeated games. Am. Econ. Rev. **95**(5), 1591–1604 (2005)

Davis, D.D., Holt, C.A.: Experimental Economics. Princeton University Press, Princeton (1993)
Fischer, G.: Contract structure, risk-sharing, and investment choice. Econometrica **81**(3), 883–939 (2013)
Fréchette, G.R., Yuksel, S.: Infinitely repeated games in the laboratory (2013)
Gneezy, U., Rustichini, A.: Pay enough or don't pay at all. Q. J. Econ. **115**(3), 791–810 (2000)
Harrison, G.W., Rutström, E.E.: Risk aversion in the laboratory. Res. Exp. Econ. **12**, 41–196 (2008)
Harrison, G.W., List, J.A., Towe, C.: Naturally occurring preferences and exogenous laboratory experiments: a case study of risk aversion. Econometrica **75**(2), 433–458 (2007)
Holt, C.A., Laury, S.K.: Risk aversion and incentive effects. Am. Econ. Rev. **92**(5), 1644–1655 (2002)
Lei, V., Noussair, C.N.: An experimental test of an optimal growth model. Am. Econ. Rev. **5**, 549–570 (2002)
Murnighan, J.K., Roth, A.E.: Expecting continued play in prisoner's dilemma games a test of several models. J. Confl. Resolut. **27**(2), 279–300 (1983)
Roth, A.E., Murnighan, J.K.: Equilibrium behavior and repeated play of the prisoner's dilemma. J. Math. Psychol. **17**(2), 189–198 (1978)
Sherstyuk, K., Tarui, N., Saijo, T.: Payment schemes in infinite-horizon experimental games. Exp. Econ. **16**(1), 125–153 (2013)
Vespa, E.: Cooperation in dynamic games: an experimental investigation. Available at SSRN 1961450 (2013)

Authors and Affiliations

Arun Gautham Chandrasekhar[1,2] · **Juan Pablo Xandri**[3]

✉ Arun Gautham Chandrasekhar
arungc@stanford.edu

✉ Juan Pablo Xandri
jxandri@uandes.cl

1 Department of Economics, Stanford University, 579 Jane Stanford Way, Stanford 94305, CA, USA

2 Palo Alto, California, USA

3 School of Business and Economics, Universidad de los Andes, Chile, Monsenor Alvaro del Portillo 12455, Las Condes, Santiago 7560026, RM, Chile