

COOPERATION WITH TIME-INCONSISTENCY

Extended Abstract for LMSC09

By Nicola Dimitri

**Professor of Economics
Faculty of Economics
University of Siena
Piazza S. Francesco 7
53100 Siena
Italy**

Dynamic games have proven to be a powerful framework to conceptualize social and economic repeated interactions, among strategic agents. Within dynamic games, repeated strategic form games have played a major role, both in the theory and in the applications (Mailath - Samuelson; 2006).

Well known achievements of such games are the celebrated “Folk Theorems”, a class of results showing that cooperative behaviour can obtain as a Nash Equilibrium in an infinitely repeated Prisoner’s Dilemma. The main intuition behind these theorems, namely that cooperation can emerge even among selfish players as long as they assign sufficiently high weights to future payoffs, lies behind incentives design in a number of real life repeated relationships.

The theory of infinitely repeated games (IRG) has been developed assuming that players are time-consistent. Consistency implies, for example, that an action planned to be taken at a future date is effectively taken when that date comes. In game and decision theory time consistency is formalized by assigning exponential weights to future payoffs.

In IRG a major implication of time consistency (exponential discounting) is that when players choose a strategy at the beginning of the game they do not need to reconsider it at later dates, as the game evolves. This is because, with exponential discounting, at any date the welfare level generated by an infinite stream of payoffs is the same, from any point in time, which implies that as the game unfolds there is no incentive to revise a strategy.

In recent years however empirical evidence showed that time consistency is often violated by individuals, as in the classical example of preference reversal (Ainslie, 2001). More specifically, preference reversal occurs when an individual facing a choice between, say, 5€ in seven days from now and 10€ in twenty days from now, he first chooses the later (higher) 10€ payoff and then, as the first payoff gets closer in time, he reverses his own choice opting for the earlier, though lower, 5€ sum.

When players, in an IRG, are time inconsistent their welfare level at various points in time along the game might change, and therefore their strategy choice in the game cannot be analyzed by simply considering the welfare, generated by the infinite stream of available payoffs, at the beginning of the game.

As diffused as time inconsistent behaviour might be, so far no attempt seems to have been made to investigate how the theory of IRG could be reformulated to take into account individuals’ inconsistency. More in general, the point that we are interested in investigating is the following: to what extent “problems” related to time consistency can be mitigated, or eliminated, when players take into proper account their own future inconsistency?

In this paper we shall provide a first contribution in this direction, by considering one of the main models proposed in the literature to formalize inconsistency, the so called Quasi-Hyperbolic

(sometimes also called Quasi-Exponential) discounting function (Phelps and Pollack, 1968; Laibson, 1997, Frederick et al (2002)).

In what follows, after having suggested how the standard IRG theory could be amended to accommodate time inconsistent players, in the case of the Infinitely Repeated Prisoner's Dilemma we compare conditions under which cooperation could emerge, when individuals are time consistent *vs* time inconsistent, where inconsistency is indeed formalized by Quasi-Hyperbolic discounting.

1. Preliminaries

In this section we set the basics of the analysis. Let $t=0,1,2,\dots,T,\dots$ be the time index and $x=(x(0),x(1),\dots,x(t),\dots)$, $y=(y(0),y(1),\dots,y(t),\dots)$ two alternative streams of payoffs available to a decision maker (DM), which for convenience we assume to be monetary amounts. Moreover, let $d(t)$, with $d(0)=1$ and $d(t)\geq d(t+1)$ be the DM discounting function. We also assume that the DM chooses at $t=0$ and can neither borrow nor lend, so that at each time $t=0,1,2$, he can get exactly what's specified initially in the payoff profile.

At time t , the payoff vector x generates the following welfare level

$$w_t(x) = \sum_{i=0}^t \frac{x(i)}{d(t-i)} + \sum_{i=t+1}^{\infty} d(i-t)x(i) \quad (1)$$

analogously, for the payoff stream y we have

$$w_t(y) = \sum_{i=0}^t \frac{y(i)}{d(t-i)} + \sum_{i=t+1}^{\infty} d(i-t)y(i) \quad (2)$$

Formulae (1)-(2) specify how, at time $t=0,1,2,\dots,T,\dots$, the individual welfare is computed by capitalizing up to date t the payoffs obtained prior to reaching it, and discounting at t the payoffs available after that date.

In repeated games it is standard to evaluate the payoff stream x , associated to a strategy profile, by considering $t=0$; namely

$$w_0(x) = \sum_{i=0}^{\infty} d(i)x(i) \quad (3)$$

This is because, for all $i=1,2,\dots$ the discounting function $d(i)$ is typically assumed to follow an exponential pattern, namely $d(i) = d(1)^i$, which implies that for all $t=0,1,2,\dots$

$$w_0(x) = \sum_{i=0}^{\infty} d(i)x(i) = d(t)w_t(x) = d(1)^t w_t(x)$$

Hence, if $w_0(x) > w_0(y)$ then, clearly, also $w_t(x) = d(1)^{-t}w_0(x) > d(1)^{-t}w_0(y) = w_t(y)$. More in general, with exponential discounting the welfare ordering of two alternative payoff streams is the same at all dates t .

This is not so with non-exponential discounting functions, so that the welfare ordering may change at different times. For example it could be that $w_0(x) > w_0(y)$ and, for some $T>0$, $w_T(x) < w_T(y)$, which would entail time inconsistency since a decision taken at $t=0$ might be revised later at $t=T$.

Therefore, if the decision maker is aware of his own being time inconsistent, how would he proceed?

2. Evaluating Welfare with Time Inconsistency

Assuming the individual to be aware of his own inconsistency, namely that at future dates his welfare ordering over two alternative payoff profile may change, we imagine that he will evaluate vector x (and analogously for vector y) by the following welfare aggregator

$$W_0(x) = W(x) = \sum_{t=0}^{\infty} d(t)w_t(x) \quad (4)$$

namely that the decision as to which payoff stream is preferable is still taken at $t=0$, like in standard game theoretic approach with exponential discounting, however based upon a discounted sum of future welfare levels, rather than simply upon payoffs.

The underlying motivation is simple. If discounting of future payoffs may lead to inconsistency, and then to a possible ambiguity for the individual on which action to take at $t=0$ then it is natural to think that the decision at $t=0$ could be taken by weighting, discounting, future welfare levels rather than payoffs.

We could observe that (4) represents a “first order” sophisticated reasoning on the part of the decision maker, who anticipating his own inconsistency discounts future welfare levels rather than payoffs. But if along the lines of (4) he would also consider, and compare, aggregators of future welfare levels at time t such as

$$W_t(x) = \sum_{i=0}^t \frac{w_i(x)}{d(t-i)} + \sum_{i=t+1}^{\infty} d(i-t)w_i(x) \quad (5)$$

then his level of sophistication would increase, making his decision more articulated since an issue of time inconsistency could also concern now the welfare aggregators $W_t(x)$.

In the next section we shall confine our attention to the order of sophistication embodied in (4), leaving to future research inconsistency issues related to (5).

In particular, we shall do so within the celebrated Prisoner’s Dilemma game, where inconsistency will be modelled by the Quasi Hyperbolic (QHD) discounting function $d(0) = 1$ and $d(t) = \beta \delta^t$, with $\beta > 0$, $0 < \delta < 1$ and $t=1,2,\dots$. More specifically we shall be interested in comparing the conditions on the parameter δ , for the cooperative outcome to obtain, in both QHD and in the Exponential Discounting (ED) model $d(t) = \delta^t$, with $t = 0,1,2,\dots$.

We think such comparison could provide some initial elements towards an understanding of whether or not awareness of one’s inconsistency could eliminate inconsistency.

3. The Infinitely Repeated Prisoner’s Dilemma

The theory of infinitely repeated games (IRG) has clarified how economic, political and social outcomes that could not be obtained as a Nash Equilibrium of the one-shot game, could instead be achieved as a Nash Equilibrium when players strategic interaction is infinitely repeated. A most notable example of this is the infinitely repeated Prisoner’s Dilemma (IRPD), such as that in Table 1 below

Table 1

	C	D
C	2,2,	0,3
D	3,0	1,1

Start considering ED, and assume (for simplicity) both players to be characterized by the same δ . Then, it is well known that if $\delta > \frac{1}{2}$ a pair of “modified trigger strategies” (Osborne, 2003) would support the cooperative outcome (2,2) at every t , namely at every repetition of the game.

We now want to discuss what would happen if both players have the same QHD function, rather than the same ED function.

3.1 Quasi Hyperbolic Discounting

Consider now (1) and (2), where $x = (2,2,2,\dots)$ is the payoff vector when the cooperative outcome realizes at every repetition, and $y = (3,1,1,\dots)$ the payoff vector when a players decides to deviate at the beginning and the one-shot-game Nash Equilibrium is played from then on.

Let's first consider the generic term of (4), for the payoff stream x

$$\begin{aligned} \beta\delta^t w_t(x) &= 2\beta\delta^t \left[\frac{1}{\beta\delta^t} + \frac{1}{\beta\delta^{t-1}} + \dots + \frac{1}{\beta\delta} + 1 + \beta\delta + \beta\delta^2 + \beta\delta^3 + \dots \right] = \\ &= 2 \left[\frac{(1-\delta^t)}{(1-\delta)} + \beta\delta^t \left(1 + \frac{\beta\delta}{(1-\delta)} \right) \right] \end{aligned}$$

Analogously

$$\begin{aligned} \beta\delta^t w_t(y) &= \beta\delta^t \left[\frac{3}{\beta\delta^t} + \frac{1}{\beta\delta^{t-1}} + \dots + \frac{1}{\beta\delta} + 1 + \beta\delta + \beta\delta^2 + \beta\delta^3 + \dots \right] = \\ &= \left[2 + \frac{(1-\delta^t)}{(1-\delta)} + \beta\delta^t \left(1 + \frac{\beta\delta}{(1-\delta)} \right) \right] \end{aligned}$$

Therefore

$$W(y) - W(x) = \lim_{T \rightarrow \infty} \sum_{t=0}^T \left[\frac{(1-2\delta)}{(1-\delta)} + \frac{\delta^t}{(1-\delta)} - \beta\delta^t \left(1 + \frac{\beta\delta}{(1-\delta)} \right) \right] = \begin{cases} \infty & \text{if } \delta < \frac{1}{2} \\ -\infty & \text{if } \delta > \frac{1}{2} \end{cases} \quad (6)$$

so that the condition for cooperation to obtain, $\delta > \frac{1}{2}$ is the same as when the discounting function is exponential, with generic term δ^t .

The above conclusion hinges upon the DM preferring to deviate at $t=0$, when he decides to do so. Such preference however depends on the value of the parameters β and δ and there would always be values of β close to one which makes immediate deviation most profitable. Conditions for cooperation would in case be independent of β .

4. Conclusions

In the paper we begun investigating how time inconsistent players, in infinitely repeated games, could take into account of their own inconsistency, if they are aware of it. We suggested that they could do so by discounting future welfare levels, rather than future payoffs. In an infinitely repeated Prisoner's Dilemma we then enquired whether or not, when players take account of their future inconsistency as modelled by Quasi Hyperbolic Discounting, conditions for cooperation differ with respect to exponential discounting, namely with respect to when they are time consistent. The analysis is a preliminary step showing that if Quasi Hyperbolic Discounters discount all future welfare levels then cooperation could emerge under the same conditions of exponential discounting. More explicitly, inconsistent and consistent players may decide to cooperate on the same parameter values. This would seem to suggest that inconsistency could be tackled by appropriately taking into account one's inconsistency. Full investigation of the issue however is left to future research.

References

Ainslie G., (2001), *Breakdown of Will*, Cambridge University Press.

Frederick S, Loewenstein G, O'Donoghue T, (2002) Time discounting and time preferences: A Critical Review, *Journal of Economic Literature* XL: 351-401.

Laibson D (1997) Golden Eggs and Hyperbolic Discounting, *Quarterly Journal of Economics* 42: 861-871.

Mailath G., Samuelson L., (2006), *Repeated Games and Reputations: Long Run Relationships*, Oxford University Press.

Osborne M., (2003), *An Introduction to Game Theory*, Oxford University Press.

Phelps, E. S., R. A. Pollak (1968), On Second-Best National Saving and Game-Equilibrium Growth, *Review of Economic Studies*, 35, 185-199.

Pollak R (1968), Consistent Planning, *Review of Economic Studies* 35: 201-208.

Strotz R (1955), Myopia and Inconsistency in Dynamic Utility Maximization, *Review of Economic Studies* 23, 165-180.