TIME INCONSISTENCY AND LEARNING IN BARGAINING GAMES

BY ZAFER AKIN

March 2005

1. The Problem and Related Literature

In our daily lives, we always face decisions to make and alternative actions to choose overtime. Traditional economic analysis expects people behave rationally (take actions maximizing their payoff or utility) and thus behave consistently (following the original contingent plan or strategy) when they make these decisions. In other words, a rational agent's goals (and strategies to achieve them) at different dates cannot be in conflict and he always agrees with his future selves. However, we always suffer from these kinds of conflicts. This is due to the vulnerability of individuals to self-deception, over-optimism, over-confidence, self-control and many other characteristics mentioned in Psychology literature. One way of incorporating some of these characteristics into the decision making analysis is to introduce time-inconsistent preferences. Hyperbolic discounting is often used in economics literature to model time-inconsistency (interchangeably, preference reversals or self-control problem) (see, Dasgupta and Maskin, 2002; Laibson, 1997; O'donoghue and Rabin, 1999).

The literature about time-inconsistency, which attracts more attention recently, mainly takes the characteristic of the agent as given and focuses on how this plays a role in different decision making problems. Some of the existing literature also mentions "time-inconsistent agents' abilities of learning to be less time-inconsistent overtime" at footnotes (see, O'donoghue and Rabin, 2001; Sarafidis, 2004). However, there is no distinct work, as far as we know, focusing on learning capabilities of time-inconsistent agents.

The rationale behind time-inconsistent learners is the following: delaying costly tasks forever, which arises as a symptom of self-control, or holding the same belief about one's self is prevented by different forces such as deadlines. One can realize his inconsistent behavior to some extent if he repeatedly fails to follow his plans and/or does not carry out actions in accordance with his beliefs. In other words, he can revise his contingent actions and update his beliefs about himself. While, generically, it is difficult to impose an evolutionary structure on time-inconsistent behavior of agents (e.g., time-inconsistent agents will vanish in the population overtime by evolutionary forces), it is possible to observe learning to be more rational in some specific strategic environments.

This paper will not address learning of time-inconsistent agents thoroughly in a general context either, but it introduces learning in a context of bargaining. We present two solution concepts formalized by Sarafidis, 2004, and their applications to bargaining games first without learning (extensions of some arguments in Akin, 2004), then we will allow learning. At each case, we examine whether there is delay in bargaining and what the equilibrium shares are. The main argument is that existence of time-inconsistent agents who may learn to be more rational overtime may explain bargaining delays to some extent.

2. Model

The existing literature on time-inconsistent preferences introduced continuum types of agents: naïve, partially naïve and sophisticated that represent different levels of unawareness of agents' self-control problems. This paper incorporates bounded rationality in the form of self-control that leads to time-inconsistency into a sequential bargaining model in which one player or both players are one of these types.

Let T = {0, 1, 2, 3...} be the discrete and infinite time space. There are two players, one of which is time-consistent Exponential Player (EA) and the other is one of three (actually four) types: time consistent Exponential Agent (EA), Naive Hyperbolic Agent (NHA), Sophisticated Hyperbolic Agent (SHA) (and partially naive agent). They are engaged in an alternating offers bargaining game. When one player is (partially) naive hyperbolic, we may have a learning motive (that is, NHA may learn from rejections to be more sophisticated in time). Let i and j represent players and t and s represent dates.

At each time t, i offer a share of size 1 pie. If j accepts the offer, the game ends and if there is rejection, then at time t+1, j offers a utility pair. If they never agree, then each player gets 0.

Players' preferences are as follows: one player is time-consistent exponential agent and he has the following sequence of discount factors: $\{1, \delta, \delta^2, \delta^3...\}$. The other player is NHA and has the following sequence of discount factors: $\{1, \beta\delta, \beta\delta^2, \beta\delta^3...\}$ where δ is the standard time-consistent impatience, β is time-inconsistent preference for immediate gratification or the self-control problem of the agent. Let β^ be a person's belief about her future self-control problems- her beliefs about what her taste for immediate gratification, β, will be in all future periods. A sophisticated person knows exactly what her future self-control problems will be in the future, therefore has perceptions β^=β. A naive person believes she will not have future self-control problems in the future, therefore has perceptions β^=1. A partially naive person has perceptions β^ between β and 1.

Three different approaches can be pursued: 1. No learning at all, 2. immediate learning and 3. Gradual learning. Without learning, β=1 and it does not evolve overtime, which means NHA believes that her self-control problem will disappear after tomorrow and she will not change this belief whatever happens in the future periods (at time t>1, she believes that she will discount t+1 by δ). In immediate learning case, we assume NHA learns immediately whenever she observes a rejection (either rejection of her offer or she rejects an offer), that is, β=1 becomes β^=β immediately after one rejection. Gradual learning examines behavior of NHA who learns to be more sophisticated gradually in time and β^ may not equal 1 and it may evolve overtime and gets closer to β. She will learn by introspection about herself and evolution of the game. We will follow Yildiz's (2004) framework to model learning.

There are two different solution concepts proposed by Sarafidis (2004), "Naive Backward Induction" (NBI) and "Equilibrium" in games with time inconsistent players (hyperbolic discounters).

In a NBI, the player (she) who has self-control problem plays a best response to what she thinks the rational opponent will play. In other words, players can rationalize what they will play during the course of the game (that is, NBI solution is rationalizable). One caveat of this concept is that since time-inconsistent player may not implement what she has planned for future and she has wrong beliefs about both herself and the rational opponent, she may be surprised by how players (including herself) play when the game proceeds. She also believes that rational agent thinks about her what she thinks about herself. That is, she is also naive about beliefs of rational agent about her. In NBI, players form beliefs about how other players will play the game by introspection and putting themselves in the shoes of other players and so on. However, this belief formation process leads NHA to anticipate opponent's action incorrectly.

In "Equilibrium", players are endowed with some beliefs about how others will play the game. Each player takes these beliefs as given and plays a best response, without questioning how and why other players have chosen to play this way (each announces their strategies in advance). In addition, since each player plays the game as others expect them to play, the original beliefs are confirmed in equilibrium. Subgame perfection is also required in "equilibrium". An unfavorable aspect of "equilibrium" is that naive time-inconsistent player may not understand why her opponent plays the way he does.

It turns out that our equilibrium concept and "equilibrium" solution concept are same. In Sarafidis (2004), NHA thinks that EA believes what she thinks about herself and this leads NHA not to understand why EA plays the way he does. What we differently assume is that second order beliefs of agents are correct. In other words, NHA knows that EA believes that she is naive. In terms of epistemological foundations of this assumption, followings can be mentioned: contemplation of NHA about her past experiences with opponent may make her to form this belief. Each making announcements of his/her strategies in advance is another alternative explanation (contemplation of NHA about what EA announces may allow her to understand why he announces the way he does). However, independent of belief formation process, both agents have beliefs about the opponent and equilibrium actions confirm these beliefs.

3. Main Results

We have results in each case where agents can learn and their types are fixed, no learning. Without learning, we showed an immediate agreement result such that being naïve pays off in the sense that the more naïve the agent is, the higher share she gets. While, generically, it is difficult to impose an evolutionary structure on time-inconsistent behavior of naïve agents (learning to be more rational); here, we further incorporate a learning model that allows naïve agents to learn as they play the game. We showed that there is a critical date at which parties reach agreement and before that date, there is no agreement. When at least one of the parties is naïve, depending on learning structure, we get delay in sequential bargaining game. Hence, existence of players who are time-inconsistent and can learn as they play the game can be another explanation for delays in bargaining.