TIME INCONSISTENCY AND LEARNING IN BARGAINING GAMES

BY ZAFER AKIN*[†]

FIRST DRAFT, DECEMBER 2004

THIS VERSION:

July 5, 2005

ABSTRACT. The literature on time-inconsistent preferences introduced naive, partially naive and sophisticated as types of agents that represent different levels of unawareness of agents' self-control problems. This paper incorporates time-inconsistent players in a sequential bargaining model. We first consider "naive" agents who never learn about their types and show that bargaining between such a player and a standard exponential agent ends in immediate agreement. The more naive a player, the higher his share. If naive agents can learn their type over time, we show that there is a critical date such that there is no agreement before that date. Hence, existence of time-inconsistent players who can learn as they play the game can be another explanation for delays in bargaining.

KEYWORDS: Hyperbolic discounting, learning, bargaining, delay

1. INTRODUCTION

In our daily lives, we always face decisions to make and alternative actions to choose over time. Traditional economic analysis expects people to behave rationally (take actions maximizing their

---

*Department of Economics, Pennsylvania State University, 608 Kern Graduate Building, University Park, PA, 16802. E-mail: zafer@psu.edu

[†] I am grateful to my advisor Prof. Kalyan Chatterjee for his guidance and advice throughout this project. I also thank participants of Wednesday meetings at Penn State. All remaining errors are mine.

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. This means he will not have a following sort of conversation with himself: "Ok, I had decided to do this before; but now, let me do something else" (e.g., "We had agreed with my coauthors that I was going to write only the introduction of our grant proposal as of tomorrow, but let me write both introduction and the literature review in one week" or "I had decided to renew my computer at least 6 months from now, but let me buy this new wide screen laptop now").

The above argument, however, misses the fact that in real life, individuals always suffer from these kinds of conflicts. This is due to the vulnerability of them to self-deception, over-optimism, over-confidence, self-control and many other characteristics mentioned in the 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 the economics literature to model time-inconsistency (see, Laibson, 1997; O'Donoghue and Rabin, 1999, 2001; Phelps and Pollak, 1968).

The literature on time-inconsistency (interchangeably, preference reversals or self-control problems) introduced naive, partially naive and sophisticated as types of agents that represent different levels of unawareness of agents' self-control problems. Naive agents are not aware of their future preference reversals at all. Sophisticated agents are fully aware of their self-control problems. Partially naive agents introduced by O'Donoghue and Rabin (2001) perceive their self-control problems to some extent.

A naive time-inconsistent agent procrastinates. However, deadlines and potentially learning prevents her from procrastinating costly tasks forever. One might argue that evolutionary "learning" would cause time-inconsistent preferences to disappear from the population. However, evolution acts over a long time horizon. Here we focus on individuals becoming less naive about their self-control problems during the course of a bargaining game.

In my earlier paper (Akin, 2004), I consider an alternating-offers bargaining game where there are time-inconsistent players who cannot learn and show an immediate agreement result by using a modified Nash equilibrium solution concept. In that paper, bargaining is the second stage of a two stage principal-agent game whose first stage is self-investment of the agent. On the other hand, in this paper, I introduce learning in a sequential bargaining context.

In the model we consider, different types of agents are engaged in an infinite horizon alternating-offers bargaining game. When we consider the interplays among different types of agents, the games involving partially naive agents are the most interesting ones. During the game, partially naive agents, by observing possible rejections, might gradually become more aware of their naivete. We model this in a similar way to Yildiz (2004), though Yildiz examines optimism about recognition process rather than self-control.

Learning works as follows: partially naive agent has an initial belief about her future self-control problems. When she observes a rejection during the course of the game, she interprets these rejections in a way that her actual self-control problems might actually be more severe than she perceives and she updates her belief accordingly. On the other side, the opponent has a trade-off between delaying the game and extracting more rent from the partially naive agent and cost of delaying (discounting). When the cost of delaying outweighs the benefit, the game ends (probably with some delay).

With time-inconsistent players, the solution concept used could be problematic. We consider two solution concepts, "naive backward induction" and "equilibrium", developed by Sarafidis (2004).

In the *naive backward induction (NBI)*, the naive player plays a best response to what she thinks the opponent will play. However, since she has wrong beliefs about herself and about what the opponent thinks about herself, she may be surprised by how players (including herself) play during the course of the game. Thus, due to this wrong belief formation, naive agent is not able to anticipate the opponent's action correctly. In the *equilibrium*, players are endowed with some beliefs about how others will play the game. Each player takes these beliefs -which turn out to be correct in equilibrium as opposed to NBI- as given and plays a best response.

We first consider naive agents who never learn about their types and we, by using "equilibrium" as the solution concept, show that bargaining between such a player and a standard exponential agent ends in immediate agreement. When we use NBI as the solution concept, there is immediate agreement when exponential agent offers first. When naive agent offers, there is one period delay. Moreover, the more naive a player, the higher his share. With NBI, we obtain perpetual disagreement between two naive agents. If naive agents can learn their type over time, we show that there is a critical date such that there is no agreement before that date. At each case where different types of agents are engaged in a sequential bargaining game, we examine whether there is delay and what the equilibrium shares are. *There are two main arguments. First one is that existence of time-inconsistent agents who may learn to be more consistent over time may explain bargaining delays to some extent and second one is that being naive makes you better off except the opponent is also naive.*

There is a growing literature about time-inconsistency. First, Strotz (1956) proposed that individuals may not have stationary preferences over time. They might value close satisfaction more than distant ones. Phelps and Pollak (1968) formalized this kind of behavior in a mathematically more convenient way, called $\beta - \delta$ approach. Afterwards, Laibson (1997) used this formalization to explain individuals' observed saving behavior. O'Donoghue and Rabin (1999a, 1999b, 2001) examined time-inconsistent individuals' decision making in different economic environments such as when to complete a task, agent-principal problems, which option to choose from a menu of options and when. What we add is to carry this formalization to the context of a non-cooperative game.

Although some of the literature just briefly talk about potential learning considerations as extensions to the existing models (see, O'Donoghue and Rabin, 2001; Sarafidis, 2004), there is no distinct work, as far as we know, focusing on this important aspect of time inconsistency. Dellavigna and Malmendier (2003) examines self-control in the market and shows that some of the observed behavior of consumers (for example in the health clubs) can be explained by the time inconsistency of the agents. They also mention the effect of learning of these agents to explain some empirical results. They say: "Consumers choosing monthly or annual contracts (out of three different contracts: pay per visit, monthly and annual) in the health clubs would on average have saved money paying per visit...In four-fifth of the cases, these contracts are terminated and learning has a large effect in this observation...It is hard to believe that individuals remain naive about their own preferences and ability after a lifetime of experience". We, in this paper, not only carry time-inconsistency to the context of non-cooperative games but we also address learning of time-inconsistent agents to be more dynamically consistent over time in this context for the first time.

We use the alternating-offers bargaining framework proposed first by Rubinstein (1982). Rubinstein assumes stationarity of preferences over time. On the other hand, Coles and Muthoo (2003) examined bargaining situations in a non-stationary environment. In their paper, they study Rubinstein's bargaining game in which the set of possible utility pairs evolves through time in a non-stationary but smooth manner. In our model, we have a non-stationary environment too. However, non-stationarity comes directly from the preferences of the players.

The rest of the paper is organized as follows. Section 2 describes the solution concepts. Section 3 introduces the formal model. Section 4 characterizes the equilibrium of the game without learning. Section 5 incorporates learning considerations into the model. Section 6 concludes the paper with a brief discussion of the results.

2. Solution Concepts

The solution concepts used here depend crucially on the beliefs agents that have about offers and about their future selves. We, therefore, discuss these beliefs first. A time-inconsistent agent may be one of three types: naive, sophisticated or partially naive. The naive hyperbolic agent (NHA) is naive about her time inconsistency, which means she thinks that she will be patient but in reality she will be impatient in future periods. The sophisticated hyperbolic agent (SHA) is fully aware of her time inconsistency and behaves accordingly. Partially naive agent is aware of his future self-control problems only to some extent[1]. Time-consistent agent is fully rational and knows opponent's type. The naive hyperbolic agent has wrong beliefs about herself and also believes that the rational agent thinks about her what she thinks about herself.

For the games including only time consistent players, we have Subgame Perfect Nash Equilibrium (SPNE) as the solution concept. For the games including only time consistent and sophisticated hyperbolic players (SHA): since SHA is fully aware of his preference reversals, he may be treated as a time consistent player with standard impatience $\beta\delta$. This implies that we have stationarity of preferences and mutually consistent beliefs. Thus, we can again apply SPNE. On the

---

[1] Types will be defined and explained formally in the next section.

other hand, If at least one of the players is NHA, then we have to define and use slightly different solution concepts. 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 a time-inconsistent player may not implement what she has planned for the 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 the rational agent thinks about her what she thinks about herself. That is, she is also naive about beliefs of the 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 the other players. However, this belief formation process leads the NHA to anticipate opponent's actions incorrectly.

In an "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 (e.g., 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. An unfavorable aspect of "equilibrium" is that the naive time-inconsistent player may not understand why her opponent plays the way he does, though she has correct beliefs about him. The following analogy can be made to understand "equilibrium" better: Agents announce their strategies in advance and each player plays a best response according to these announcements. When we talk about announcements, incredible threats arise as an issue. To prevent incredible threats, we impose subgame perfection. Since for a naive agent, what she plans to do in the future may differ from the actual actions in the future, the "equilibrium" requires both the planned and the actual strategies of the naive agent to be best responses to the strategies of the opponent's strategies.

We now construct necessary notation for formal definitions. We will consider a two player alternating-offers bargaining game. One player is time consistent (EA) and the other is a time-inconsistent naive agent (NHA) [2]. We will create an infinite sequence of fictitious players from the NHA. Let NHA_t represent the $t-period$ self of the NHA. NHA_t has available actions that the NHA has from period t onwards.

Let s_{EA} be a history dependent strategy for the EA and s_{NHA_t} be a history dependent strategy of NHA_t or

$$s_{NHA_t} = \{s^k_{NHA_t}\}^{\infty}_{k=t}$$

Define strategy s_{NHA} as the sequence of moves, each of which is the move of NHA_t at time t (first component of s_{NHA_t}) $\forall t$ or

$$s_{NHA} = \{s^t_{NHA_t}\}^{\infty}_{t=0}$$

The above arguments motivate the following formal definition of the solution concepts.

Definition 1. *A strategy profile* $s = (s_{EA}, s_{NHA_0}, s_{NHA_1}, s_{NHA_2}, ...)$ *is an "equilibrium" if:*

1. *Strategy* s_{EA} *is a best response to* s_{NHA} *for the* EA*,*
2. *Strategy* s_{NHA_t} *is a best response to* s_{EA} *for each* NHA_t*,*
3. *In all subgames,* s *induces an equilibrium (satisfies condition 1 and 2).*

---

[2] As mentioned, these solution concepts are mainly for games including at least one naive hyperbolic agent. Games with other types will also be examined.

Definition 2. *A strategy profile* $s = (s_{EA}, s_{NHA_0}, s_{NHA_1}, s_{NHA_2}, ...)$ *constitutes "Naive Backwards induction" (NBI) solution if:*

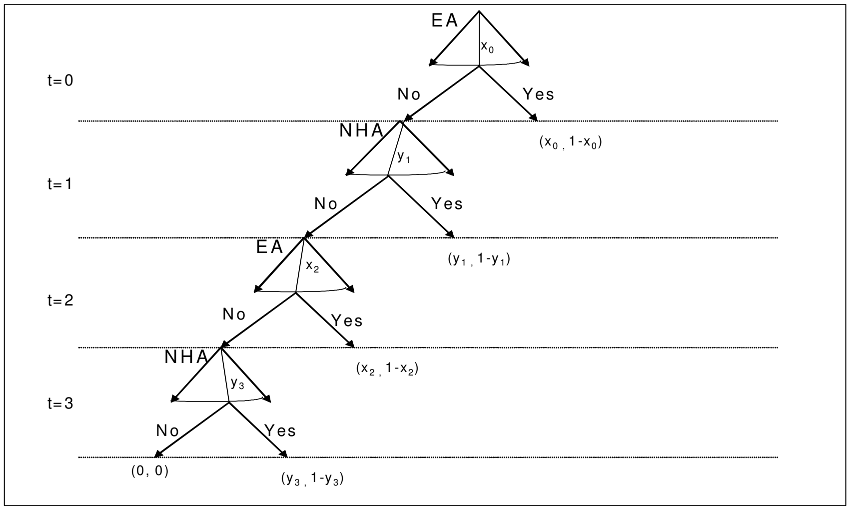
1. Strategy s_{EA} *is a best response to* s_{NHA} *for the* EA,

2. Each strategy s_{NHA_t} *survives backwards induction in the game between the* EA *and the* NHA_t.

In order to make the two solution concepts and understand the difference between them, we will give an example (see the game-tree). Let a NHA (she) and an EA (he) play a finite sequential bargaining game to share a size 1 pie. The EA will offer at $t = 0$ and the game ends at $t = 3$. The pie will vanish at the end of the third period if they cannot reach an agreement. As shown in the figure, the EA offers $(x_0, 1 - x_0)$, where x_0 is his share, and the NHA accepts of rejects. If she accepts, the pie is allocated according to the offer $(x_0, 1 - x_0)$. If she rejects, she makes a counter offer $(y_1, 1 - y_1)$ at $t = 1$, where y_1 is the share of the EA. If the game proceeds to second period, the EA offers $(x_2, 1 - x_2)$. If the game proceeds to $t = 3$, the NHA offers $(y_3, 1 - y_3)$ and if the EA accepts, the pie is allocated, if he does not, they both get zero.

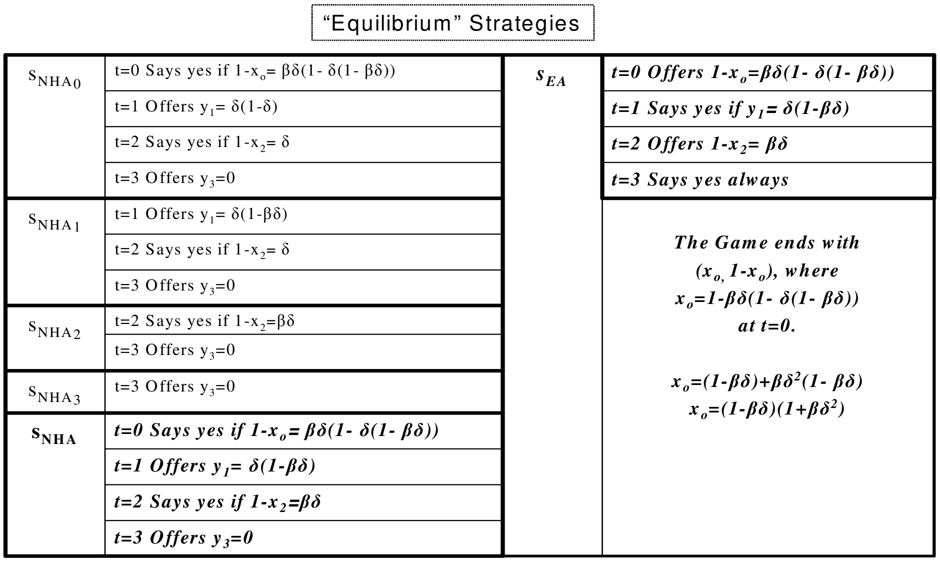
The tables show the "equilibrium" and "Naive Backward Induction" solution strategies of each player. We now will argue that these strategies actually constitute "equilibrium" and "Naive Backward Induction" solutions, respectively.

In order for a strategy profile to constitute an "equilibrium", there are three conditions (see definition 1). The equilibrium strategies below constitute an equilibrium since they satisfy those conditions. For the first condition, s_{EA} has to be a best response to s_{NHA} for the EA. We can see this easily because each action in s_{EA} in different periods is a best response to each action in s_{NHA} and the EA cannot increase his payoff by deviating. Each s_{NHA_t} is a best response to s_{EA}



Game-tree for the example

Figure 1:



“Equilibrium” Strategies

s_{NHA0}	t=0 Says yes if $1\text{-}x_o= \beta\delta(1\text{-} \delta(1\text{-} \beta\delta))$	s_{EA}	***t=0 Offers $1\text{-}x_o=\beta\delta(1\text{-} \delta(1\text{-} \beta\delta))$***
	t=1 Offers $y_1= \delta(1\text{-}\delta)$		***t=1 Says yes if $y_1= \delta(1\text{-}\beta\delta)$***
	t=2 Says yes if $1\text{-}x_2= \delta$		***t=2 Offers $1\text{-}x_2= \beta\delta$***
	t=3 Offers y_3=0		***t=3 Says yes always***
s_{NHA1}	t=1 Offers $y_1= \delta(1\text{-}\beta\delta)$		***The Game ends with $(x_o, 1\text{-}x_o)$, where $x_o=1\text{-}\beta\delta(1\text{-} \delta(1\text{-} \beta\delta))$ at t=0. $x_o=(1\text{-}\beta\delta)+\beta\delta^2(1\text{-} \beta\delta)$ $x_o=(1\text{-}\beta\delta)(1+\beta\delta^2)$***
	t=2 Says yes if $1\text{-}x_2= \delta$		
	t=3 Offers y_3=0		
s_{NHA2}	t=2 Says yes if $1\text{-}x_2=\beta\delta$		
	t=3 Offers y_3=0		
s_{NHA3}	t=3 Offers y_3=0		
$\mathbf{s_{NHA}}$	***t=0 Says yes if $1\text{-}x_o= \beta\delta(1\text{-} \delta(1\text{-} \beta\delta))$***		
	t=1 Offers $y_1= \delta(1\text{-}\beta\delta)$		
	t=2 Says yes if $1\text{-}x_2=\beta\delta$		
	t=3 Offers y_3=0		

Figure 2:

“Naive Backward Induction” Strategies

s_{NHA0}	t=0 Says yes if $1-x_o=\beta\delta(1-\delta(1-\delta))$	s_{EA}	***t=0 Offers*** $1-x_o=\beta\delta(1-\delta(1-\delta))$
	t=1 Offers $y_1=\delta(1-\delta)$		***t=1 Says yes if*** $y_1=\delta(1-\beta\delta)$
	t=2 Says yes if $1-x_2=\delta$		***t=2 Offers*** $1-x_2=\beta\delta$
	t=3 Offers $y_3=0$		***t=3 Says yes always***
s_{NHA1}	t=1 Offers $y_1=\delta(1-\delta)$		***The Game ends with*** $(x_o, 1-x_o)$, ***where*** $x_o=1-\beta\delta(1-\delta(1-\delta))$ ***at t=0.*** $x_o=1-\beta\delta(1+(-\delta)+(-\delta)^2)$
	t=2 Says yes if $1-x_2=\delta$		
	t=3 Offers $y_3=0$		
s_{NHA2}	t=2 Says yes if $1-x_2=\beta\delta$		
	t=3 Offers $y_3=0$		
s_{NHA3}	t=3 Offers $y_3=0$		
$\mathbf{s_{NHA}}$	***t=0 Says yes if*** $1-x_o=\beta\delta(1-\delta(1-\delta))$		
	t=1 Offers $y_1=\delta(1-\delta)$		
	t=2 Says yes if $1-x_2=\delta$		
	t=3 Offers $y_3=0$		

Figure 3:

for each NHA_t. To see this, check s_{NHA_0} first. If the NHA_0 rejects the offer of EA at $t = 0$, then she thinks she will offer $y_1 = \delta(1 - \delta)$ at $t = 1$ and knows that EA will reject this. At $t = 2$, if she accepts the EA's offer, she gets $\beta\delta$ (which has a discounted value of $\beta\delta^2\beta\delta$ at $t = 0$), if she rejects, she will get 1 for sure (which has a discounted value of $\beta\delta^3$ at $t = 0$). She compares all these possibilities from her perspective[3]. It turns out that accepting the EA's offer at $t = 0$ is the optimal strategy for NHA_0. For NHA_1, it can easily be seen that s_{NHA_1} is the optimal strategy and so on. For the third condition, we can repeat this analysis for each subgame and see that the first two conditions are satisfied in all subgames.

In order for a strategy profile to constitute a "NBI", there are two conditions (see definition 2). The NBI strategies above constitute a NBI solution since they satisfy those conditions. Again, s_{EA} has to be a best response to s_{NHA} for the EA. We now check whether the EA can increase his payoff by offering less to the NHA_0. If he offers less, he knows she will make an offer that he will certainly reject and he will get $(1 - \beta\delta)$ for sure at $t = 2$. However, the discounted value of this payoff is always smaller than what he can get from offering according to his original strategy $[1 - \beta\delta(1 - \delta(1 - \delta)) > \delta^2(1 - \beta\delta)]$. Thus, s_{EA} is a best response to s_{NHA}. The second condition is also satisfied as follows: From the perspective of each NHA, her strategy has to survive backwards induction between her and the EA. For 0^{th} self of NHA, she thinks as follows: "I can get 1 at $t = 3$ for sure. To convince me, the EA will offer δ at $t = 2$. To convince him, I have to offer at least $\delta(1 - \delta)$ at $t = 1$ and at $t = 0$, he will offer me $\beta\delta(1 - \delta(1 - \delta))$ that I will accept". Similar arguments can be made for each self of NHA. Thus, each NHA_t survives backwards induction.

On the other hand, if there are two naive agents are engaged in a game, then we have to revise the above definitions as follows:

---

[3] specifically: $\beta\delta(1 - \delta(1 - \beta\delta)) > \beta\delta^3 \ or \ \frac{1-\delta}{1-\beta} > \delta^2$

$NHA(i)$: i^{th} Naive Hyperbolic Agent. EA: Exponential Agent (Rational). $NHA_t(i) = t-$ *period* self of the $NHA(i)$.

Players are represented by ”i”and”j” where $i \neq j \in \{1, 2\}$. Now s represents a strategy profile for these players. Necessary notation for the new definition is as follows:

$$s = \{s_{NHA}(1); s_{NHA}(2)\}$$

$$= \{s_{NHA_t}(1), s_{NHA_{t+1}}(1), s_{NHA_{t+2}}(1), ...; s_{NHA_t}(2), s_{NHA_{t+1}}(2), s_{NHA_{t+2}}(2), ...\}$$

$$= \{\underbrace{[s^t_{NHA_t}(1), s^{t+1}_{NHA_t}(1), s^{t+2}_{NHA_t}(1)...]}_{s_{NHA_t}(1)}, \underbrace{[s^{t+1}_{NHA_{t+1}}(1), s^{t+2}_{NHA_{t+1}}(1), ...]}_{s_{NHA_{t+1}}(1)}, ...;$$
$$\underbrace{[s^t_{NHA_t}(2), s^{t+1}_{NHA_t}(2), s^{t+2}_{NHA_t}(2)...]}_{s_{NHA_t}(2)}, \underbrace{[s^{t+1}_{NHA_{t+1}}(2), s^{t+2}_{NHA_{t+1}}(2), ...]}_{s_{NHA_{t+1}}(2)}, ...\}$$

$$s_{NHA}(i) = \{s^t_{NHA_t}(i), s^{t+1}_{NHA_{t+1}}(i), s^{t+2}_{NHA_{t+2}}(i)...\} = \{s^k_{NHA_k}(i)\}_{k=t}^{\infty}.$$

Definition 3. *A strategy profile s constitutes an "equilibrium" in the existence of two naive agents if*

1. *Each $s_{NHA_t}(i)$ is a best response to $s_{NHA}(j)$ for every NHA_t.*
2. *In all subgames, s induces an equilibrium (satisfies condition 1).*

Informally, each self of the naive player plays a best response to what will actually be played by the other naive player. We also impose subgame perfection.

The following figure shows a simple example where two naive agents are playing a game. $NHA(1)$ thinks at t=1 that she will choose "8 at t = 3" over "7 at t = 2" and "10 at t = 3" over "9 at t = 2" . However, at t = 2, she prefers "7 at t = 2" to "8 at t = 3" and "9 at t = 2" to "10 at t = 3" because of her tendency for immediate gratification. Similarly, $NHA(2)$ thinks at t=1 that he will choose "6 at t = 3" over "5 at t = 2" and "4 at t = 3" over "3 at t = 2". However, at t = 2, he prefers "5 at t = 2" to "6 at t = 3" and "3 at t = 2" to "4 at t = 3". In other words,

$$\begin{aligned} at\ t &= 1,\ NHA(1) \rightarrow 8_{t=3} \succ 7_{t=2}\ and\ 10_{t=3} \succ 9_{t=2} \\ at\ t &= 2,\ NHA(1) \rightarrow 8_{t=3} \prec 7_{t=2}\ and\ 10_{t=3} \prec 9_{t=2} \end{aligned}$$

and

$$\begin{aligned} at\ t &= 1,\ NHA(2) \rightarrow 6_{t=3} \succ 5_{t=2}\ and\ 4_{t=3} \succ 3_{t=2} \\ at\ t &= 2,\ NHA(2) \rightarrow 6_{t=3} \prec 5_{t=2}\ and\ 4_{t=3} \prec 3_{t=2} \end{aligned}$$

We can summarize beliefs of each agent as follows:

$NHA(1):$

She believes that she will play "right" independent of which node she is at.

She believes that he will play "Right".

$NHA(2):$

He believes that he will play "Left".

He believes that she will play "left" independent of which node she is at.

In the example, if $x = 9$, then ((out, r, r); Left) constitutes an equilibrium. If $x = 7$, then ((in, r, r); Left) constitutes an equilibrium. We can see this as follows.

$x = 9:$

$NHA(1)$ thinks that $NHA(2)$ will play "Right" if she plays "in" although he announces his strategy as "Left". Then, she will play "right" and get 8 at $t = 3$. Since playing "out" gives 9, she plays "out".

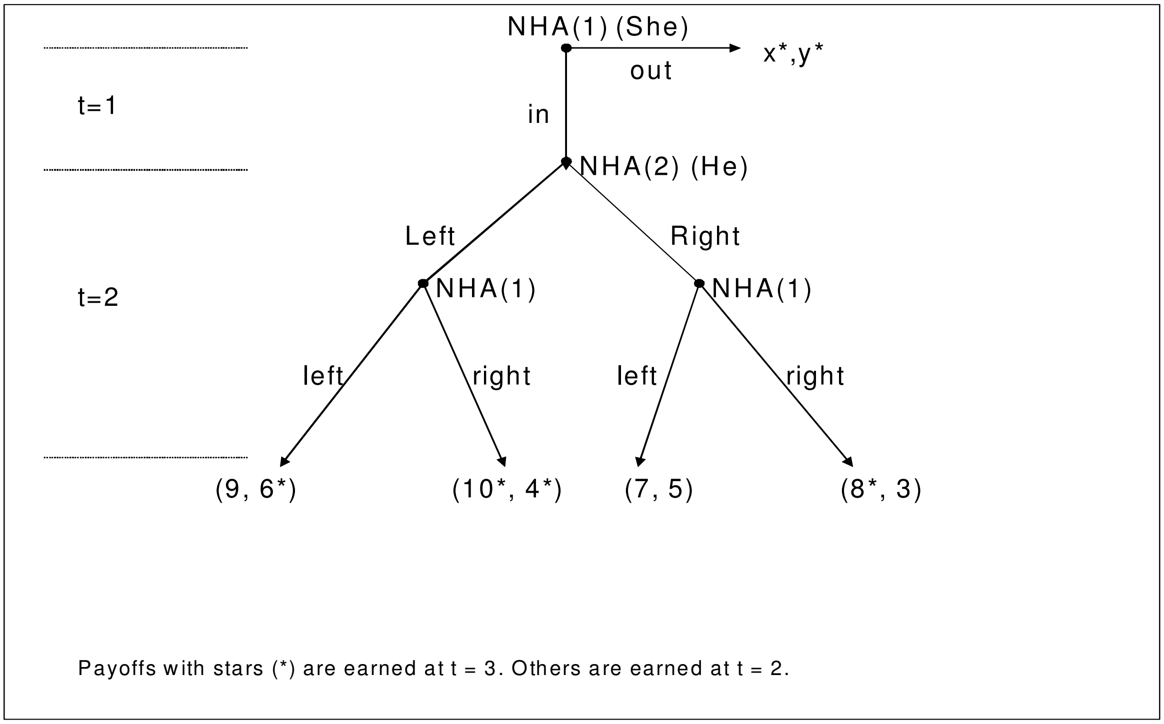


Figure 4: Game-tree for the example. This is a slightly different version of the example of Sarafidis (2004).

On the other hand, $NHA(2)$ thinks that $NHA(1)$ will play "left" regardless of where she is although she announces his strategy as "right". Then, he will play "Left" and get 6 at $t = 3$ instead of playing "Right" and getting 5 at $t = 2$.

So, NHA_1 plays "out" and the game ends.

$x = 7:$

$NHA(1)$ thinks that $NHA(2)$ will play "Right" if she plays "in" although he announces his strategy as "Left". Then, she will play "right" and get 8 at $t = 3$. Since playing "out" gives 7, she plays "in".

On the other hand, $NHA(2)$ thinks that $NHA(1)$ will play "left" regardless of at which node she is although she announces his strategy as "right". Then, he will play "Left" and get 6 at $t = 3$ instead of playing "Right" and getting 5.

What actually happens at $t = 2$ is that they both change their minds and play the other strategies that they have. Specifically, NHA_1 plays "left" at both her information sets, which confirms the original belief of NHA_2 about her and NHA_2 plays "Right" confirming the original belief of the NHA_1 about him.

So, NHA_1 plays "in", NHA_2 plays "Right" and NHA_1 plays "left".

Thus, each of the naive agents plays a best response to what the other player will actually play and beliefs of each agent about the other is confirmed in equilibrium but not the beliefs about themselves. Obviously, the announced strategies of each agent are vulnerable to changes because of their self-control problems.

Now, we will introduce the model and examine what the equilibrium outcome will be in the cases where firstly, we assume no updating of beliefs and secondly, we allow updating of beliefs.

3. MODEL

Let $T = \{0, 1, 2, 3, ...\}$ denote the infinite set of possible agreement times. Let $i \neq j \in \{1, 2\}$ and $t, s \in T$ represent players and dates respectively. Let U be the set of feasible utility pairs, $U = \left\{u \in [0,1]^2 \,|u^1 + u^2 \leq 1\right\}$. At each time t, i offers a utility pair $u = \left\{u^1, u^2\right\} \in U$. 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 can be one of four types: time-consistent exponential type (EA), Naive type (NHA), Sophisticated type (SHA) and partially naive type. The EA has the following sequence of discount factors: $\left\{1, \delta, \delta^2, \delta^3, ...\right\}$. The NHA and the SHA have the following sequence of discount factors: $\left\{1, \beta\delta, \beta\delta^2, \beta\delta^3, ...\right\}$ and the Partially naive agent has the following sequence of discount factors: $\left\{1, \beta\delta, \widehat{\beta}\delta^2, \widehat{\beta}\delta^3, ...\right\}$where δ is the standard time-consistent impatience with $\delta \in (0, 1)$, β is time-inconsistent preference for immediate gratification or the self-control problem of the agent with $\beta \in (0, 1)$. Let $\widehat{\beta}$ 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. The NHA believes she will not have future self-control problems in the future, therefore has perceptions $\widehat{\beta} = 1$. The SHA knows exactly what her future self-control problems will be in the future, therefore has perceptions $\widehat{\beta} = \beta$. The partially naive person has perceptions $\widehat{\beta} \in (\beta, 1)$.

An evolutionary preference structure can be imposed on naive and partially naive agents' beliefs because they are not fully aware of their future preference reversals. We, therefore, can incorporate learning into the environments where there are naive and partially naive players. Three different learning approaches can be pursued: 1. *No learning at all,* 2. *Immediate learning* and 3. *Gradual learning. Without learning,* $\widehat{\beta} \in (\beta, 1]$ does not evolve over time, which means agents believe that their self-control problem, β, will disappear (β will be $\widehat{\beta} = 1$) or diminish (β will be $\widehat{\beta}$) after tomorrow and they will not change this belief whatever happens in the future periods (at time $t > 1$, she believes that she will discount $t+1$ by $\widehat{\beta}\delta$). In *immediate learning* case, we assume players learn immediately whenever they observe a rejection (either rejection of their offer or they reject an offer), that is, $1 \geq \widehat{\beta} > \beta$ becomes $\widehat{\beta} = \beta$ immediately after one rejection. *Gradual learning* examines behavior of players who learn to be more sophisticated gradually in time, that is, $\widehat{\beta}$ may not equal 1 and may evolve over time and gets closer to β. They learn by introspection about themselves during the evolution of the game. We will follow Yildiz's (2004) framework to model learning.

We now will examine these three approaches in order. First, we assume that there is no learning at all. Then, we will examine the other two approaches allowing learning.

4. Equilibrium Without Learning

We will study "Equilibrium" as the first solution concept. We will show that without learning, players will immediately agree. Unique solution directy follows from Rubinstein. Afterwards, we will examine NBI as the second solution concept.

We have two solution concepts and in different economics environments one can be more plausible to use. Especially, if there is no precedent or past experience between players, NBI is the appropriate solution concept to use because the only possible way to form beliefs about the opponent is to put themselves in the shoes of other players. Since basic assumptions of NBI fits better to our framework, we will focus on it more.

Since it will be needed in the following results, it is useful to write down the equilibrium of the Rubinstein bargaining game where players are exponential type and have different discount factors δ_1 and δ_2. We can write the result as either the limit case of the finite horizon game or the recursive way of solving it like in Shaked-Sutton, using stationarity of the game.

Remark 1. *In the infinite horizon alternating-offers game with both players have exponential discounting with discount factors* δ_1 *and* δ_2*, the equilibrium payoffs are:*

$$(x^*, 1 - x^*) \text{ where } x^* = \frac{1 - \delta_2}{1 - \delta_1 \delta_2}$$

and x^* *is the payoff of the agent* 1*, making first offer.*

Informally, a Nash equilibrium involves players playing best responses to their beliefs about the other player and the beliefs are correct and mutually consistent. Not surprisingly, with time-inconsistent players, the last requirement is difficult to satisfy. Given this caveat, subgame perfectness is defined in the usual way as being Nash after every history.

4.1. "Equilibrium". We know from the general framework of the Rubinstein model that players will reach an agreement immediately in equilibrium since there is discounting. First player offers an x^* such that second player accepts and realized payoffs would be $(x^*, 1 - x^*)$. A NHA is relatively impatient for tomorrow and thinks that she will be more patient for periods onward. If future periods are reached then she will reoptimise, thereby choosing actions that are different from the ones foreseen. In SHA case, the rational opponent (he) will anticipate that the SHA is sophisticated and that if future periods are reached, her discount rate will be $\beta\delta$ and she is aware of this. Thus, he makes offers according to the conjecture that the SHA's discounting rate is effectively $\beta\delta$. Following claim shows that SHA will get the amount of payoff where she has an effective discount rate $'\beta\delta'$ and EA has $'\delta'$.

Claim 1. *In the infinite horizon alternating-offers game on a size* 1 *pie with only one player (sophisticated) has hyperbolic discounting, SHA, the equilibrium payoffs are the following:*

if SHA makes the first offer, payoffs are $(x^*, 1-x^*)$ *where* x^* *is the share of SHA and equals to* $\frac{1-\delta}{1-\beta\delta^2}$.

if EA makes the first offer, payoffs are $(y^*, 1-y^*)$ *where* y^* *is the share of EA and equals to* $\beta\delta\frac{1-\beta\delta}{1-\beta\delta^2}$.

Proof. We can use the Shaked-Sutton approach (1984) to find the equilibrium payoffs in this problem. In order to use this approach, we need to have stationarity. Stationarity (in the sense of Rubinstein, 1982) can be interpreted as, the preference of (x, t) over $(y, t+1)$ is independent of t. This means that the preference over getting x at time t and getting y at time $t+1$ is independent of time where x and y are the shares of a size 1 pie. In other words, the agents do not have any preference reversals over time.

As it is explained above, the game, with an EA and a SHA, turns out to be the Rubinstein alternating-offers bargaining game with stationary preferences $\delta_1 = \beta\delta$ and $\delta_2 = \delta$. Then, the following would be true since we have stationarity:

$$1 - x^* = \delta(1 - y^*)$$

$$y^* = \beta\delta x^*$$

which gives the following results:

$$y^* = \beta\delta(\frac{1-\delta}{1-\beta\delta^2}) \textit{ and } x^* = \frac{1-\delta}{1-\beta\delta^2}$$

So, when SHA offers first then she (SHA) gets x^* and when EA offers first, she (SHA) will get $1 - y^*$. ■

As we can infer from the discussion in the solution concept section, in the concept of "equilibrium", the agents are exogeneously given some beliefs and they play best responses to these beliefs. Strategies that are best responses to these beliefs have to satisfy the three conditions in the definition of the equilibrium. According to the definition and the example, we can conclude that when a naive agent and a rational agent play an alternating-offers bargaining game, the rational agent plays as if he plays against a sophisticated hyperbolic agent or a rational agent who has $\beta\delta$ discount factor. This is optimal strategy from the perspective of him because the naive agent takes these beliefs as given and she plays a best response to these beliefs. She cannot understand why he announces his strategies in this way but she plays a best response anyway. She does think that these may be incredible threats but actually even if she thinks like that, the announced strategies induce an equilibrium at each subgame because if the subgames are actually reached, she will again be impatient and these beliefs will be confirmed.

Claim 2. *In the infinite horizon alternating-offers game with only one player (naive) has hyperbolic discounting, NHA, the equilibrium payoffs are like the following:*

1.*If the EA makes the first offer, payoffs are:*

$$(x^*, 1-x^*) \text{ where } x^* \text{is the share of the EA and equals to } \frac{1-\beta\delta}{1-\beta\delta^2}.$$

2.*If the NHA makes the first offer, payoffs are:*

$$(x^*, 1-x^*) \text{ where } x^* \text{is the share of the NHA and equals to } \frac{1-\delta}{1-\beta\delta^2}.$$

Proof. 1. The EA, by offering x^*, he should make the NHA indifferent between accepting and rejecting. He knows that the NHA has discount factor $\beta\delta$ between $t = 0$ and $t = 1$, so he should make an offer x^* satisfying

$$1 - x^* = \beta\delta(1 - y^*)$$

Since the EA knows that the NHA will discount any future period with $\beta\delta$, he offers as if he is playing against a rational agent with discount factor $\beta\delta$. As a best response, the NHA offers y^* to the EA satisfying

$$y^* = \delta x^*$$

we get the result in the claim. This result actually is the result where two rational agents are playing the game and one of them has δ (offering first), the other has $\beta\delta$ discount factor.

2. This is very similar to the above case that only the order of offers changes. The NHA, by offering x^*, she should make the EA indifferent between accepting and rejecting. So, she should make an offer x^* satisfying

$$1 - x^* = \delta(1 - y^*)$$

Since the EA knows that the NHA will discount any future period with $\beta\delta$, he offers as if he is playing against a rational agent with discount factor $\beta\delta$. So, he offers y^* to the NHA satisfying

$$y^* = \beta\delta x^*$$

we get the result in the claim. ■

It turns out that the EA offers same shares to both the sophisticated agent and the naive agent. The rationale behind the latter is the following: the EA plays a best response to what will actually be played by the NHA. Although, the NHA is not aware that she may follow different strategies in the future and she does not agree with the EA, she plays a best response to what she is given (the announced strategies of the EA). This leads to the above result even the future periods are not reached and even the EA's strategies seem as incredible threats. These strategies are actually the subgame perfect strategies because if the future periods would have been reached, then the original beliefs of each agent would be confirmed.

We now check the payoffs of the agents when they have different characteristics. We assume complete information that means each player knows the other player's characteristic, e.g., in a

NHA-NHA game, a naive player knows that the other player is naive but she does not know she, herself, behaves naively.

In the table below, row player, player 1, makes the first offer to the column player, player 2, in the alternating-offers bargaining game. The payoffs in the table are such that first entry is the payoff of the row player and second entry is the payoff of the column player, e.g., P_{23} represents the game where SHA makes the first offer to the NHA and gets the first entry of the P_{23}. P_{11} is the case of the classical alternating-offers bargaining game. P_{12} and P_{21} are the results of the *claim* 1. P_{13} and P_{31} are the results of the *claim* 2. P_{22} is the same case where two EA with effective discount factors $'\beta\delta'$. P_{23} is a similar case to the P_{13} with EA has an effective discount factor $'\beta\delta'$. P_{32} is a similar case to the P_{31} with EA has an effective discount factor $'\beta\delta'$. P_{33} is a similar case to the P_{22} becauseeach thinks that she will be patient after tomorrow and that the other is naive but since they will take the other's beliefs as given, it will give same result with P_{22}.

We can compare the payoffs of each player by checking the equilibrium payoff table. The table implies the following theorem:

Theorem 1. *Assume each agent has the same time-consistent impatience, δ. Further, assume that the self-control problems of the agents, if any, are also same, β and $\delta > \beta$. If we use "equilibrium" as the solution concept, then regardless of whether a player offers first or second and regardless of the opponent's type (EA, NHA or SHA), his payoff will be $P_{EA} > P_{NHA} = P_{SHA}$ according to his type.*

Proof. We can infer the result from the table below:

Equilibrium (given beliefs-confirmed, subgame perfection)

		Player 2		
		EA	SHA	NHA
Player 1	EA	$1/(1+\delta), \delta/(1+\delta)$	$(1-\beta\delta)/(1-\beta\delta^2)$, $\beta\delta(1-\delta)/(1-\beta\delta^2)$	$(1-\beta\delta)/(1-\beta\delta^2)$, $\beta\delta(1-\delta)/(1-\beta\delta^2)$
	SHA	$(1-\delta)/(1-\beta\delta^2)$, $1-(1-\delta)/(1-\beta\delta^2)$	$1/(1+\beta\delta), \beta\delta/(1+\beta\delta)$	$1/(1+\beta\delta), \beta\delta/(1+\beta\delta)$
	NHA	$(1-\delta)/(1-\beta\delta^2)$, $1-(1-\delta)/(1-\beta\delta^2)$	$1/(1+\beta\delta), \beta\delta/(1+\beta\delta)$	$1/(1+\beta\delta), \beta\delta/(1+\beta\delta)$

Figure 5:

1. If a player makes the first offer, then we can show why $P_{EA} > P_{NHA} = P_{SHA}$ is satisfied as follows:

$$\frac{1}{1+\delta} > \frac{1-\delta}{1-\beta\delta^2}; \frac{1-\beta\delta}{1-\beta\delta^2} > \frac{1}{1+\beta\delta}$$

2. If a player makes the second offer, then in terms of payoffs, if she is EA, then she does better than the case where she is NHA or SHA, in other words, $P_{EA} > P_{NHA} = P_{SHA}$, since

$$\frac{\delta}{1+\delta} > \frac{\beta\delta(1-\delta)}{1-\beta\delta^2}; 1-\frac{1-\delta}{1-\beta\delta^2} > \frac{\beta\delta}{1+\beta\delta}$$

■

Briefly, this theorem states that given the assumptions on preferences, regardless of your opponent's type, you get a higher payoff if you are an EA and less payoff if you are a NHA or a SHA. Thus, rational players always do better than hyperbolic agents.

4.2. Naive Backwards Induction (NBI). In this part, we will study bargaining games including at least one naive agent by using the NBI as the solution concept. We will check each case separately. The assumption is that NHA never updates her beliefs about herself and the opponent.

Proposition 1. *If two naive agents are engaged in an alternating-offers bargaining game, then NBI gives "never agree" as the only solution.*

Proof. The above result seems odd because "never agree" gives the worst payoff to each of the players so they must not follow this strategy. This arises because neither of the naive agents is aware that the game will evolve like this (that they end up with zero payoff). Moreover, since there is no learning, they do not update their beliefs and change their strategies (they stay naive however the game evolves).

Since the belief structure is crucial to understand the proof, we will mention it again. Naive agent believes that she will be time-consistent from tomorrow on. She also believes that the opponent thinks about her what she thinks about herself. In other words, she believes that the opponent believes that she will be time-consistent from tomorrow on. That is, she is naive about herself and also what the opponent thinks about herself, too.

Yildiz (2003) points out that if it is common knowledge that the players will remain sufficiently optimistic for a sufficiently long future, then in equilibrium, they will agree immediately. That is, excessive optimism alone cannot be a reason for a delay in agreement. In our case, players stay optimistic about their own preferences forever, but we do not get an immediate aggreement, even an agreement. Optimism in our context is the degree of naivete. Completely naive agent is the most optimistic agent because she is so optimistic that she believes that she will be time-consistent from tomorrow on for sure (optimism is connected to time-consistency because it is advantageous in terms of payoff and implies no self-control problem).

Since there is informational (and behavioral) deficiencies of agents about themselves, here, we cannot talk about common knowledge of optimism or its persistency over time. This informational structure makes the above result possible. Now, we examine what happens during the course of the game.

Let i offers $(i, j) = (x_t, 1 - x_t)$ at each even time period t $(t = 0, 2, 4...)$ and j offers $(i, j) = (y_t, 1 - y_t)$ at each odd time period t $(t = 1, 3, 5...)$. The following notation for beliefs will be used: at time t, let

x_s^i denote what i thinks she (i) will offer at any future even time period $s > t$,

x_s^j denote what j thinks i will offer at any future even time period $s > t$,

y_s^j denote what j thinks she (j) will offer at any future odd time period $s > t$,

y_s^i denote what i thinks j will offer at any future odd time period $s > t$.

Note that with time-inconsistent agents, $y_s^j \neq y_s^i$ (s even) and $x_s^j \neq x_s^i$ (s odd) are possible beliefs. Also note that $y_s^j \neq y_s^i \neq y_s$ (s even) and $x_s^j \neq x_s^i \neq x_s$ (s odd) are possible situations.

Since agents have same discount factors δ and β, we only need to examine the very first period $t = 0$. This means the offers will be symmetric (same) and will not change over time (since there is no learning). This implies $x_s = x_{s+2}$ for every $s = 2, 4, 6...$ and $y_s = y_{s+2}$ for every $s = 1, 3, 5...$

At $t = 0$, i offers $(i, j) = (x_0, 1 - x_0)$ satisfying:

$$1 - x_0 = \beta\delta(1 - y_1^i)$$

i thinks that $y_1 = \frac{\delta}{1+\delta}$ (or $y_1^i = \frac{\delta}{1+\delta}$) because she believes she will be time-consistent from tomorrow on and she also believes that j has the same belief. Thus, i offers $(i, j) = (x_0, 1 - x_0) = (1 - \frac{\beta\delta}{1+\delta}, \frac{\beta\delta}{1+\delta})$.

However, j knows that i will discount tomorrow and the next day by $\beta\delta$ as opposed to her ($i's$) beliefs. At $t = 0$, j thinks her offer at $t = 1$ will satisfy the following:

$$y_1^j = \beta\delta x_2^j$$

j thinks that she will be time-consistent after tomorrow and she also thinks that i thinks in the same way. Moreover, i will think that she (i) will be time-consistent at $t = 2$, so j believes $x_2^i = x_2 = \frac{1}{1+\delta}$. This means if j rejects $i's$ offer at $t = 0$, she believes, at $t = 1$, she can get $1 - y_1^j = 1 - \frac{\beta\delta}{1+\delta}$ which is worth $\beta\delta(1 - y_1^j) = \beta\delta(1 - \frac{\beta\delta}{1+\delta})$ at $t = 0$. Since $\beta\delta(1 - \frac{\beta\delta}{1+\delta}) > \frac{\beta\delta}{1+\delta}$, she rejects $i's$ offer at $t = 0$.

When $t = 1$ comes, j offers $y_1 = \frac{\beta\delta}{1+\delta}$ but i rejects this with the hope of getting $x_2^i = 1 - \frac{\beta\delta}{1+\delta}$ at $t = 2$ (since $\beta\delta(1 - \frac{\beta\delta}{1+\delta}) > \frac{\beta\delta}{1+\delta}$). Since there is no learning, this cycle continues forever and they will not agree and they both end up with zero payoff. ■

Introducing learning in the above situation is plausible because after some time, players naturally realize that they are naive to some extend and their optimism is lowered making them consent to a lower share. The conjecture with learning is that (see section 5) there is a critical date that parties will not agree up to that date of which they probably are not aware of and agree on that date.

Proposition 2. *Let an EA (he) and a NHA (she) play the alternating-offers bargaining game, then NBI gives the following result:*

1. *If the EA is the first proposer, then equilibrium shares are* $(x^*, 1 - x^*) = (1 - \frac{\beta\delta}{1+\delta}, \frac{\beta\delta}{1+\delta})$ *where* x^* *is the share of the EA,*

2. *If the NHA is the first proposer, then she offers* $(x, 1 - x) = (\frac{1}{1+\delta}, \frac{\delta}{1+\delta})$ *where* x *is the share of the NHA, the EA rejects this offer at* $t = 0$ *and he offers as in the previous case and the game*

ends with those shares at $t = 1$.

Proof. 1. If the EA is the first proposer and offers $1 - x^* = \frac{\beta\delta}{1+\delta}$ to the NHA, since this is the highest payoff that she can get, she accepts it. She will reject any offer less than this because she anticipates to get $1 - y_1^{NHA} = \frac{1}{1+\delta}$ tomorrow if she rejects today. From the perspective of the EA, this is also optimal because there is no learning and he cannot expect a higher share in the future because of this.

2. When the NHA is the first proposer, since she thinks the EA has the same beliefs with her about herself, she thinks the EA will offer $y_1^{NHA} = \frac{1}{1+\delta}$. To convince him, she offers him $1-x = \frac{\delta}{1+\delta}$ at $t = 0$. However, he can get $1 - \frac{\beta\delta}{1+\delta}$ at $t = 1$ if he rejects her offer which is worth $\delta(1 - \frac{\beta\delta}{1+\delta})$ at $t = 0$. Since $\delta(1 - \frac{\beta\delta}{1+\delta}) > \frac{\delta}{1+\delta}$, he rejects her offer at $t = 0$, offers $y_1 = \frac{\beta\delta}{1+\delta}$ at $t = 1$ and she accepts this. ■

Proposition 3. *Let a SHA (he) and a NHA (she) play the alternating-offers bargaining game, then NBI gives the following result:*

1. *If the SHA is the first proposer, then equilibrium shares are* $(x^*, 1-x^*) = (1-\beta\delta\frac{1-\beta\delta}{1-\beta\delta^2}, \beta\delta\frac{1-\beta\delta}{1-\beta\delta^2})$ *where* x^* *is the share of the SHA.*

2. *If the NHA is the first proposer, then she offers* $(x^*, 1-x^*) = (1-\beta\delta\frac{1-\delta}{1-\beta\delta^2}, \beta\delta\frac{1-\delta}{1-\beta\delta^2})$ *where* x^* *is the share of the NHA. The SHA rejects this offer at* $t = 0$ *and he offers as in the previous case and the game ends with those shares at* $t = 1$.

Proof. 1. When the SHA offers first, he offers the NHA what she thinks she can get at most tomorrow if she rejects today. If he offers less, then she will reject this offer. To convince her, he offers $1 - x = \beta\delta(1 - y)$ where $(1 - y) = \frac{1-\beta\delta}{1-\beta\delta^2}$ ($\frac{1-\beta\delta}{1-\beta\delta^2}$ is the highest share that she can get if she rejects today, check first remark).

2. The NHA believes that she will be time-consistent after tomorrow and that the SHA believes the same thing about her. In order to convince him, she will offer x satisfying:

$$1 - x = \beta\delta(1 - y)$$

Due to her wrong belief about him that he can get at most $1 - y = \frac{1-\delta}{1-\beta\delta^2}$ in the next period, she offers $\beta\delta\frac{1-\delta}{1-\beta\delta^2}$ to him. However, the SHA can get $1 - \beta\delta\frac{1-\beta\delta}{1-\beta\delta^2}$ by waiting (rejecting) one period. He makes the following comparison:

$$\underbrace{\beta\delta \qquad \underbrace{(1 - \beta\delta\frac{1-\beta\delta}{1-\beta\delta^2})}_{\text{reject the NHA's offer and get this next period}}}_{\text{payoff from rejection}} \quad > \quad \underbrace{\beta\delta\frac{1-\delta}{1-\beta\delta^2}}_{\text{payoff from acceptance}}$$

$$1 - \beta\delta^2 - \beta\delta + \beta^2\delta^2 > 1 - \delta \Rightarrow 1 > \beta$$

Thus, at $t = 0$, the SHA rejects the NHA's offer, at $t = 1$ offers $\beta\delta\frac{1-\beta\delta}{1-\beta\delta^2}$ to her and the game ends with these shares. ■

The table above shows the payoffs of different types of agents in an alternating-offers bargaining game where we apply NBI as the solution concept.

We now can compare the payoffs of each player. The table implies the following theorem:

Theorem 2. *Assume each agent has the same time-consistent impatience, δ. Further, assume that the self-control problems of the agents, if any, are also same, β and $\delta > \beta$. Then, the second proposer will always does better in case where she is NHA than the case where she is SHA (except the opponent is also naive). The more naive she is, the higher share she gets (This is valid for the first proposer case for some specific parameter values).*

Naive Backwards Induction

		Player 2		
		EA	SHA	NHA
Player 1	EA	$1/(1+\delta)$, $\delta/(1+\delta)$	$(1-\beta\delta)/(1-\beta\delta^2)$, $\beta\delta(1-\delta)/(1-\beta\delta^2)$	$1-\beta\delta/(1+\delta)$, $\beta\delta/(1+\delta)$
	SHA	$(1-\delta)/(1-\beta\delta^2)$, $1-(1-\delta)/(1-\beta\delta^2)$	$1/(1+\beta\delta)$, $\beta\delta/(1+\beta\delta)$	$1-\beta\delta(1-\beta\delta)/(1-\beta\delta^2)$, $\beta\delta(1-\beta\delta)/(1-\beta\delta^2)$
	NHA	* $\beta\delta/(1+\delta)$, $1-\beta\delta/(1+\delta)$	* $\beta\delta(1-\beta\delta)/(1-\beta\delta^2)$, $1-\beta\delta(1-\beta\delta)/(1-\beta\delta^2)$	** 0, 0

* These payoffs are obtained with one period delay (at t=1, not at t=0).

** This represents the perpetual disagreement.

Figure 6:

Proof. If a player makes the second offer, then in terms of payoffs, if she is EA, then she does better than the case where she is NHA who does better than the case where she is SHA (except the opponent is also naive where they both end up with nothing). In other words, $P_{EA} > P_{NHA} > P_{SHA}$, since

$$\frac{\delta}{1+\delta} > \frac{\beta\delta}{1+\delta} > \frac{\beta\delta(1-\delta)}{1-\beta\delta^2}; 1 - \frac{1-\delta}{1-\beta\delta^2} > \frac{\beta\delta - \beta^2\delta^2}{1-\beta\delta^2} > \frac{\beta\delta}{1+\beta\delta}$$

If a player makes the first offer, then, the above result is satisfied for some specific parameter values. When the naive agent is the first proposer there is always delay. When the opponent is EA or SHA, it is only one period but when the opponent is also naive, there is actually no agreement. The following conditions have to be satisfied:

$$1.\ \beta\delta \underbrace{\frac{\beta\delta}{1+\delta}}_{\text{Earned with one period delay}} > \frac{1-\delta}{1-\beta\delta^2} \quad and \quad 2.\ \beta\delta \underbrace{\frac{\beta\delta(1-\beta\delta)}{1-\beta\delta^2}}_{\text{Earned with one period delay}} > \frac{1}{1+\beta\delta}$$

but since condition 2 implies condition 1, only condition 2 has to be satisfied. For those parameters satisfying condition 2, the theorem holds for the first proposer. ■

5. Equilibrium With Learning: Unconsciously Waiting To Be Persuaded

We will now study the other two approaches, namely immediate and gradual learning by using NBI as the solution concept. Since learning is an issue for only naive and partially naive agents, from this point on, at least one of the parties engaging in bargaining is naive or partially naive. The reason that we obtain an immediate agreement result without learning is the persistence of the NHA in being naive (persistent optimism about herself to be time-consistent in the future). No matter how the game proceeds, she believes that she will be time-consistent after today. More importantly, since principal believes that naive agent will stay as naive whatever happens, he gives up, in some sense, and offers a share that confirms the NHA's wrong belief about herself. He cannot do anything but accept her as she is, because he does not have any better option.

Yildiz (2003) examines the effect of optimism about making offers in the future in a sequential bargaining model and he shows that if players will remain sufficiently optimistic for a sufficiently long future, then in equilibrium, they will agree immediately. He also states that "the players may have differing beliefs about the discount rates. As in the case of bargaining breakdown, this will not yield any delay in equilibrium, provided that the players do not update their beliefs about the future discount rates as they play the game". This is what we showed in the previous section that as long as there is no updating in beliefs about the future discount rates, there will be an immediate agreement. Our case is the case where excessive and persistent optimism does not cause delay in bargaining.

We will now allow players to learn as they play the game. If we let the EA hold optimistic beliefs about the naive player that she (NHA) may change her beliefs about herself (basically waiting for the NHA to be more sophisticated in time), then since the EA gets more from sophisticated agent, it may be optimal for him to wait. If he holds this optimistic belief, then he makes offers, which will be rejected by the NHA, to make her realize her self-control problem. The EA, given his beliefs, decides what to do based on the trade-off between cost of waiting and getting higher expected share. The NHA will be in a situation where *"she will unconsciously wait to be persuaded"*.

As stated in O'Donoghue and Rabin (2001), delaying forever or sticking to same belief about herself is prevented by different forces such as deadlines and learning. After repeatedly planning to do a task in the near future or holding the same belief about herself, not carrying out these plans or not acting in accordance with those beliefs, the person may realize the uselessness of such plans or beliefs and may just do the task now or update her belief about herself. While generalization of this learning process is questionable in real-life, we observe such learning in specific strategic environments.

Our learning approach is parallel with Yildiz (2004). In that paper, he allows players being optimistic about their bargaining power (measured as the probability of making offers), but they can also learn as they play the game. In our context, we allow the rational agent to hold optimistic belief about the NHA and the NHA to update her beliefs in time. Of course, she is, naturally, not aware of this characteristic of herself (that is she is not aware that she will update her beliefs). In fact, this unawareness may lead to delay in bargaining. Since the EA is optimistic about her belief updating, in order to make her update her beliefs, he makes offers that are not expected and that will be rejected by the NHA. A sequence of rejected offers will make the NHA realize her self-control problems and become more sophisticated.

5.1. Immediate Learning.

Assume that when there is one rejection, the NHA immediately realize her time-inconsistency (she becomes sophisticated).

Proposition 4. 1. *Let the EA offer first. If* $\beta\delta(1-\beta\delta^2) > (1-\delta^2)(1+\delta-\beta\delta^2)$ *then, there will be one period delay. If* $\beta\delta(1-\beta\delta^2) \leq (1-\delta^2)(1+\delta-\beta\delta^2)$ *then, there will be no delay.*

2. *Let the NHA offer first. Then, there will always be one-period delay.*

Proof. 1. If the EA offers less than $x = 1-\frac{\beta\delta}{1+\delta}$, then the NHA will reject it by thinking that next period she can get at least this much share of the pie, so he can get at at most $1-\frac{\beta\delta}{1+\delta}$ today. If he makes a rejected offer, then she immediately becomes a SHA. Then she offers $1-\frac{1-\delta}{1-\beta\delta^2}$ to the EA (from the previous part). Then, the trade-off of the EA is deciding between getting $1-\frac{\beta\delta}{1+\delta}$ today and getting $1-\frac{1-\delta}{1-\beta\delta^2}$ tomorrow: If

$$1-\frac{\beta\delta}{1+\delta} < \delta(1-\frac{1-\delta}{1-\beta\delta^2})$$

or

$$(1-\delta^2)(1+\delta-\beta\delta^2) < \beta\delta(1-\beta\delta^2)$$

then, there will be agreement at $t=1$ (not at $t=0$) with $y = 1 - \frac{1-\delta}{1-\beta\delta^2}$.

2. The NHA, at $t=0$, offers $1-y = \delta - \frac{\beta\delta^2}{1+\delta}$ to the EA. If he rejects this, she becomes a SHA and the EA offers $x = \frac{\beta\delta(1-\delta)}{1-\beta\delta^2}$ and get $1-x = \frac{1-\beta\delta}{1-\beta\delta^2}$. Then, the trade-off of the EA is deciding between getting $1-y = \delta - \frac{\beta\delta^2}{1+\delta}$ today and getting $1-x = \frac{1-\beta\delta}{1-\beta\delta^2}$ tomorrow. However, since the following is always satisfied,

$$\begin{aligned}
\delta - \frac{\beta\delta^2}{1+\delta} &< \delta\frac{1-\beta\delta}{1-\beta\delta^2} \\
(1+\delta-\beta\delta)(1-\beta\delta^2) &< (1-\beta\delta)(1+\delta) \\
\beta^2\delta^3 &< \beta\delta^3
\end{aligned}$$

there will always be one period delay. So, the EA offers $x = \frac{\beta\delta(1-\delta)}{1-\beta\delta^2}$ to the NHA at $t=1$ and she accepts it. ■

5.2. Gradual Learning. To make the learning process clear, following examples can be given:

Example 1. *Think about a student who decides to buy a pass for the gym in school to attend it regularly (to lose weight or for bodybuilding). He has optimistic belief that he will keep going to the gym regularly. However, it turns out that he is not able to do this due to different reasons (other activities, boredom, laziness). At the beginning of the next semester, he will again consider to buy the pass, but since he has this experience from the past (in other words, he realizes his self-control problem to some extent), he makes different commitments (buy it with his friend so they go together) or (he realized enough such that)he does not buy it at all. However, he continues to have the same experience again and again, eventually he gives up buying gym pass.*

Example 2. *Think about a student who decides to audit (no registration, just sit in a class and listen) different courses not from his major (e.g., to increase his job opportunities) (or registers morning sections of core courses). He has optimistic belief that he will continue to audit these courses till the end of the semester (he will wake up early and attend classes regularly) but he is not able to do this due to different reasons (other courses, boredom, laziness). At the beginning of the next semester, he will again consider to audit some courses (register morning sections), but since he has this experience from the past, he audits fewer courses than before (registers less morning sections). However, he continues to have the same experience again and again, eventually he gives up auditing any courses (registering any morning section).*

These are some examples that include interactions among the agent's selves. It would not be true to generalize this in the sense that naive (or partially naive) agent will be sophisticated in all situations that she faces. However, in particular environments, naive agents may actually learn and update their beliefs about themselves. In strategic environments including interaction with different players, which is more challenging and requires more careful thinking, this learning process may tend to be faster or to occur earlier during the play.

We assume that partially naive agent holds an initial belief (probability of using discount factor $\beta\delta$ in the future) about her future self-control problem. She does not think that she will update her belief but when she faces with rejections, she will actually update her beliefs. We assume the partially naive agent's beliefs have beta distributions that are widely used in statistical learning models. The learning model closely follows Yildiz (2004).

We will fix any positive integers $m_{\beta\delta}$ and n with $1 \leq m_{\beta\delta} \leq n-2$ where n measures firmness of the partially naive agent's prior belief. We assume that her initial belief, $t = 0$, that she will also be impatient in the future is $\frac{m_{\beta\delta}}{n}$; after $t = m$ rejections, her belief becomes $\frac{m_{\beta\delta}+m}{n+m}$ (and $1-\frac{m_{\beta\delta}+m}{n+m}$ is the probability that she will use δ as her discount factor in the future) at any date $s \geq t = m$. This updating structure arises when the agent believes that using $\beta\delta$ discount factor in the future is a random variable distributed with some unknown parameter α that measures the probability of the agent using $\beta\delta$ at any date t, and α is distributed with a beta distribution with paramaters $m_{\beta\delta}$ and n.

What we described in the last paragraph is not common knowledge. Interestingly, naivete leads to the following situation that the partially naive agent is not aware of this updating structure but we assume the rational agent knows this and actually, acquisition of this knowledge is the reason for a possible delay.

The argument above implies that the perceived discount factor of the NHA at any date $s \geq t = m$ will be

$$\delta_m = \widehat{\beta}\delta = [\frac{m_{\beta\delta}+m}{n+m}\beta + (1-\frac{m_{\beta\delta}+m}{n+m})]\delta = [1-\frac{m_{\beta\delta}+m}{n+m}(1-\beta)]\delta$$

This implies that the partially naive agent's perception of her β is fixed and equal to $\widehat{\beta}$ for the entire future. In other words, her belief about date 50 at $t = 49$, which depends highly on the history, will be quite different than his belief about date 50 at $t = 0$. What the intrinsic assumption here is that she is not perfectly "forward looking". She does not take into account that she may change her beliefs in the future too as she is doing now (when she observes a rejection). Being partially naive here means that each rejection is unexpected. Yildiz (2004) also has a similar assumption (e.g., agents' beliefs at date 0 about dates 100 and 200 are identical).

Note that perceived discount factor of the naive agent, δ_t, is inversely related to the number of rejections, m :

$$\delta > \delta_0 > \delta_1 > \delta_2 > ... > \delta_m$$

When the number of rejections goes to infinity, the partially naive agent becomes sophisticated (as $m \to \infty$, $\delta_m \to \beta\delta$). Under this learning scheme, since n represents the firmness of initial belief of the partially naive agent, a person is perfectly naive if $1 \leq m_{\beta\delta} << n \to \infty$. In other words, a larger n implies a more severe naivete. In addition, learning slows down in time with this specification. For the sake of learning argument being significant, first, we assume the agent is not so firm in her initial belief. In the case where the agent is too firm, likelihood of obtaining delay is low as we will mention a result that relates the firmness level with the extent of bargaining delay.

In this game, under this kind of learning and information acquisition assumptions, the EA is the only one who determines the outcome of the game. The partially naive agent plays the game as she is supposed to play and the EA knows this. Depending on the partially naive agent's learning process, the EA specifies the resulting shares and the time at which the game finishes. Following lemma states one aspect of this feature of the game:

Lemma 1. The NHA's offers are never accepted by the EA.

Proof. At any time $t = 2k+1$, $k = 0, 1, 2...$, the partially naive agent makes an offer, which we call as $y_{2k+1}(\delta_{2k+1})$ and the EA's offers at even periods as $x_{2k}(\delta_{2k})$. If we reach at time $t = 2k+1$, this means $m = 2k+1$ rejections occurred as of time t. She has been updating her beliefs during this time and now, her perception on discount factors of herself for any time $t > m = 2k+1$ is δ_{2k+1}. Note that at $t = 2k+1$, she discounts payoffs at $t = 2k+2$ again by $\beta\delta$ but thinks she will discount payoffs at each $t > 2k+2$ by δ_{2k+1}.

From previous arguments, she offers the following at odd periods:

$$y_{2k+1}(\delta_{2k+1}) = \delta(\frac{1-\delta_{2k+1}}{1-\delta\delta_{2k+1}})$$

We also know that, the EA offers

$$x_{2k}(\delta_{2k}) = 1 - \beta\delta(\frac{1-\delta}{1-\delta\delta_{2k}})$$

at any even period $t = 2k$, $k = 0, 1, 2$...Now we would like to show $y_{2k+1}(\delta_{2k+1}) < \delta x_{2k+2}(\delta_{2k+2})$.

$$\delta(\frac{1-\delta_{2k+1}}{1-\delta\delta_{2k+1}}) < \delta(1 - \beta\delta(\frac{1-\delta}{1-\delta\delta_{2k+2}}))$$

$$1 > \beta\delta(\frac{1-\delta}{1-\delta\delta_{2k+2}}) + \frac{1-\delta_{2k+1}}{1-\delta\delta_{2k+1}}$$

$$1 > \beta\delta(\frac{1-\delta}{1-\delta\delta_{2k+2}}) + \frac{1-\delta_{2k+2}}{1-\delta\delta_{2k+2}} > \beta\delta(\frac{1-\delta}{1-\delta\delta_{2k+2}}) + \frac{1-\delta_{2k+1}}{1-\delta\delta_{2k+1}}$$

If the first inequality is satisfied, we are done because the second inequality is already satisfied.

$$\begin{aligned} 1 &> \beta\delta(\frac{1-\delta}{1-\delta\delta_{2k+2}}) + \frac{1-\delta_{2k+2}}{1-\delta\delta_{2k+2}} \\ &\Rightarrow 1-\delta\delta_{2k+2} > 1-\delta_{2k+2} + \beta\delta - \beta\delta^2 \\ &\Rightarrow \delta_{2k+2} > \beta\delta \end{aligned}$$

The last inequality is true by definition, so this implies $y_{2k+1}(\delta_{2k+1}) < \delta x_{2k+2}(\delta_{2k+2})$ $\forall k$. Thus, EA rejects $y_{2k+1}(\delta_{2k+1})$ $\forall k$. ■

The rationale behind this result is that the partially naive agent always offers the present value of what she thinks her rational opponent expects to earn from rejecting. However, since she will update her beliefs in case of a rejection, his expectation is always higher than what she predicts. This is because by rejecting the offer, he makes her a little more sophisticated that allows him to extract more from her at the next period and she is not aware of this.

The partially naive player is not perfectly firm ($\frac{m_{\beta\delta}}{n} > 0$) that she updates her belief as she is rejected but this updating process slows down over time. The following theorem shows that one sided learning in the existence of partially naive agents may explain delays in bargaining games. Basic intution behind this is that the rational agent can extract more share from the partially naive agent by making her more sophisticated by rejecting her offers and delaying the game but since delay is costly, when the cost of delaying exceeds this benefit, he finishes the game. Thus, this trade off between benefit and cost of delaying motivates the following theorem.

Theorem 3. In the sequential bargaining game between a partially naive player and a rational player, there exists a t^* such that before t^* players do not reach an agreement and at each time $t \geq t^*$ when the rational agent offers, players reach an agreement immediately.

Proof. By Lemma 1, we know that the partially naive agent's offers will be rejected at each odd period. Then, the rational agent will compare following payoffs that he can get at each even period:

$$x_0(\delta_0), \delta^2 x_2(\delta_2), ..., \delta^{2k-2} x_{2k-2}(\delta_{2k-2}), \delta^{2k} x_{2k}(\delta_{2k}), \delta^{2k+2} x_{2k+2}(\delta_{2k+2}), ...$$

Since he is time consistent and will make the same comparison at any given period of time, he will choose the largest element of the above sequence today and will implement what he decided today.

Note that payoff that he can get is increasing in time but waiting is costly. Hence, there should be an optimal waiting time that allows him to extract as highest share as possible from the partially naive agent. Define k^* as follows:

$$k^* = \arg\max_k \{\{\delta^{2k} x_{2k}(\delta_{2k})\}_{k=0}^{\infty}\}$$

Since at $t^* = 2k^*$, it is not optimal to delay the game anymore, he offers $(x_{t^*}(\delta_{t^*}), 1 - x_{t^*}(\delta_{t*}))$, she accepts and the game ends at $t = t^*$. We have this treshold t^* because as time passes, learning slows down and the additional payoff that he expects by waiting due to increase in sophistication of the partially naive agent is offset by the loss of waiting. In any case of multiplicity of optimal value of k, we take the minimum of those k values as k^*. ■

In brief, theorem 3 states that if we let the rational agent hold optimistic beliefs about the partially naive agent and let her learn in time (she is unaware of this), then depending on the parameter values, delay in bargaining may occur. It may well be the case that $t^* = k^* = 0$. The following corollary shows that if the prior belief of the partially naive agent is sufficiently low (n is sufficiently high) then, there will be an immediate agreement (e.g., $t^* = k^* = 0$).

Corollary 1. For any given β and δ, there exists some n^* such that $\forall n \geq n^*$, the players reach an agreement immediately in case that the rational player is first proposer and if the partially naive agent is offering first, then there will be one period delay by Lemma 1.

Proof. Given any β and δ, in order for players to reach an immediate agreement, the following condition must be satisfied $\forall k$:

$$x_{2k}(\delta_{2k}) \geq \delta^2 x_{2k+2}(\delta_{2k+2}) or$$

$$\frac{x_{2k}(\delta_{2k})}{x_{2k+2}(\delta_{2k+2})} \geq \delta^2$$

Define the following function:

$$F(\delta;\beta,n,m_{\beta\delta},k)=\delta^2 x_{2k+2}(\delta_{2k+2})-x_{2k}(\delta_{2k})$$

$$\begin{aligned} \textit{where } x_{2k}(\delta_{2k}) &= 1-\frac{\beta\delta(1-\delta)}{1-\delta\delta_{2k}} \textit{ and} \\ \delta_{2k} &= [1-(1-\beta)(\frac{m_{\beta\delta}+2k}{n+2k})]\delta \end{aligned}$$

Note that $\forall k$, $\delta_{2k}\rightarrow\delta$ and $\frac{x_{2k}(\delta_{2k})}{x_{2k+2}(\delta_{2k+2})}\rightarrow 1$ as $n\rightarrow\infty$. Also, $F(\delta;\beta,n,m_{\beta\delta},k)<0$ implies immediate agreement. As $n\rightarrow\infty$, $F(\delta;\beta,n,m_{\beta\delta},k)\rightarrow\delta^2-1<0$. Thus, there exists an n^* such that for any $n\geq n^*$, the above condition is satisfied. This is the case where the rational agent is the first proposer. If the partially naive agent is the first proposer, then by *Lemma* 1, for any $n\geq n^*$ the rational agent rejects her offer at $t=0$ and by the above argument, at $t=1$, players immediately agree. ■

Figure 1 shows delay depending on "δ" and some specific parameter values $\beta=0.5$, $m_{\beta\delta}=1$ and $k=0$. As the corollary suggests, for any given patience level, δ, we can find a firmness level of the agent, n, that will lead to immediate agreement in the bargaining game ($F(\delta;\beta,n,m_{\beta\delta},k)\leq 0$, e.g., given $\beta=0.5$, $m_{\beta\delta}=1$ and $\delta=0.995$, for all values of n satisfying $n\geq n^*=3050$, we get immediate agreement result).

The following corollary shows that as long as rational agent is patient enough, there will not be an immediate agreement.

Corollary 2. For any given $m_{\beta\delta}$, β and n satisfying condition 1 below, there exists a δ^* such that for every $\delta\geq\delta^*$ there will be $2k+2$, $k=0,1,2...$, periods delay (or $m=2k+2$ rejections) in case that the rational agent is the first proposer and if the partially naive agent is offering first, then there will be $2k+3$ period delay (or $m = 2k+3$ rejections) by Lemma 1.

$$n \geq \lceil m_{\beta\delta} + ((m_{\beta\delta}+2k)(m_{\beta\delta}+2k+2)(\frac{1-\beta}{\beta}))\rceil \qquad \text{(Condition 1)}$$

where $\lceil x \rceil$ is the ceiling function that gives *the smallest integer greater than or equal to* x.

Proof. Given any $m_{\beta\delta}$, β and n, if the following condition is satistied then, there will be $2k+2$ periods delay:

$$\begin{aligned} \delta^2 x_{2k+2}(\delta_{2k+2}) &\geq x_{2k}(\delta_{2k}) \ \mathit{or} & (2) \\ \delta^2 &\geq \frac{x_{2k}(\delta_{2k})}{x_{2k+2}(\delta_{2k+2})} \\ \mathit{where}\ x_{2k}(\delta_{2k}) &= 1 - \frac{\beta\delta(1-\delta)}{1-\delta\delta_{2k}} \ \mathit{and}\ \delta_{2k} = [1-(1-\beta)(\frac{m_{\beta\delta}+2k}{n+2k})]\delta \end{aligned}$$

Define function $F(\delta;\beta,n,m_{\beta\delta},k)$ as in the proof of $Corollary$ 1:

$$F(\delta;\beta,n,m_{\beta\delta},k) = \delta^2 x_{2k+2}(\delta_{2k+2}) - x_{2k}(\delta_{2k})$$

$$F(\delta;\beta,n,m_{\beta\delta},k) = \delta^2 - \frac{\beta\delta^3(1-\delta)}{1-\delta^2[1-(1-\beta)(\frac{m_{\beta\delta}+2k+2}{n+2k+2})]} - 1 + \frac{\beta\delta(1-\delta)}{1-\delta^2[1-(1-\beta)(\frac{m_{\beta\delta}+2k}{n+2k})]}$$

Then, (2) implies that for any given $\beta, m_{\beta\delta}, n$, if there exists a δ^* such that for any $\delta \geq \delta^*$

$$F(\delta;\beta,n,m_{\beta\delta},k) > 0 \qquad (3)$$

is satisfied then, there will be $2k+2$ periods delay.

It is easy to show that $F(\delta;\beta,n,m_{\beta\delta},k)$ is continuous in δ and $\lim_{\delta\to 1} F(\delta;\beta,n,m_{\beta\delta},k) = 0$ for every parameter value of $\beta, n, m_{\beta\delta}$ and k. Note that if $\frac{dF(\delta;\beta,n,m_{\beta\delta},k)}{d\delta}|_{\delta=1} < 0$, then (3) is true. In other words, when δ gets close to one, $F(\delta;\beta,n,m_{\beta\delta},k)$ approaches zero from first quadrant or $F(\delta;\beta,n,m_{\beta\delta},k) > 0$ for values of δ close to 1. Now it is enough to show that

$$\frac{dF(\delta;\beta,n,m_{\beta\delta},k)}{d\delta}|_{\delta=1} < 0$$

to prove the corollary.

$$F(\delta;\beta,n,m_{\beta\delta},k)=\delta^2-\frac{\beta\delta^3(1-\delta)}{1-\delta^2[1-(1-\beta)(\frac{m_{\beta\delta}+2k+2}{n+2k+2})]}-1+\frac{\beta\delta(1-\delta)}{1-\delta^2[1-(1-\beta)(\frac{m_{\beta\delta}+2k}{n+2k})]}$$

$$\frac{dF(\delta;\beta,n,m_{\beta\delta},k)}{d\delta}=2\delta$$

$$-\frac{\beta\delta^2(3-4\delta)(1-\delta^2[1-(1-\beta)(\frac{m_{\beta\delta}+2k+2}{n+2k+2})])-\beta\delta^4(1-\delta)(-2+2(1-\beta)(\frac{m_{\beta\delta}+2k+2}{n+2k+2}))}{(1-\delta^2[1-(1-\beta)(\frac{m_{\beta\delta}+2k+2}{n+2k+2})])^2}$$

$$+\frac{\beta(1-2\delta)(1-\delta^2[1-(1-\beta)(\frac{m_{\beta\delta}+2k}{n+2k})])-\beta\delta^2(1-\delta)(-2+2(1-\beta)(\frac{m_{\beta\delta}+2k}{n+2k}))}{(1-\delta^2[1-(1-\beta)(\frac{m_{\beta\delta}+2k}{n+2k})])^2}$$

$$\frac{dF(\delta;\beta,n,m_{\beta\delta},k)}{d\delta}|_{\delta=1}<0\Rightarrow 2+\frac{\beta}{(1-\beta)(\frac{m_{\beta\delta}+2k+2}{n+2k+2})}-\frac{\beta}{(1-\beta)(\frac{m_{\beta\delta}+2k}{n+2k})}<0$$

$$\Rightarrow n>m_{\beta\delta}+((m_{\beta\delta}+2k)(m_{\beta\delta}+2k+2)(\frac{1-\beta}{\beta})) \qquad \text{(Condition 1)}$$

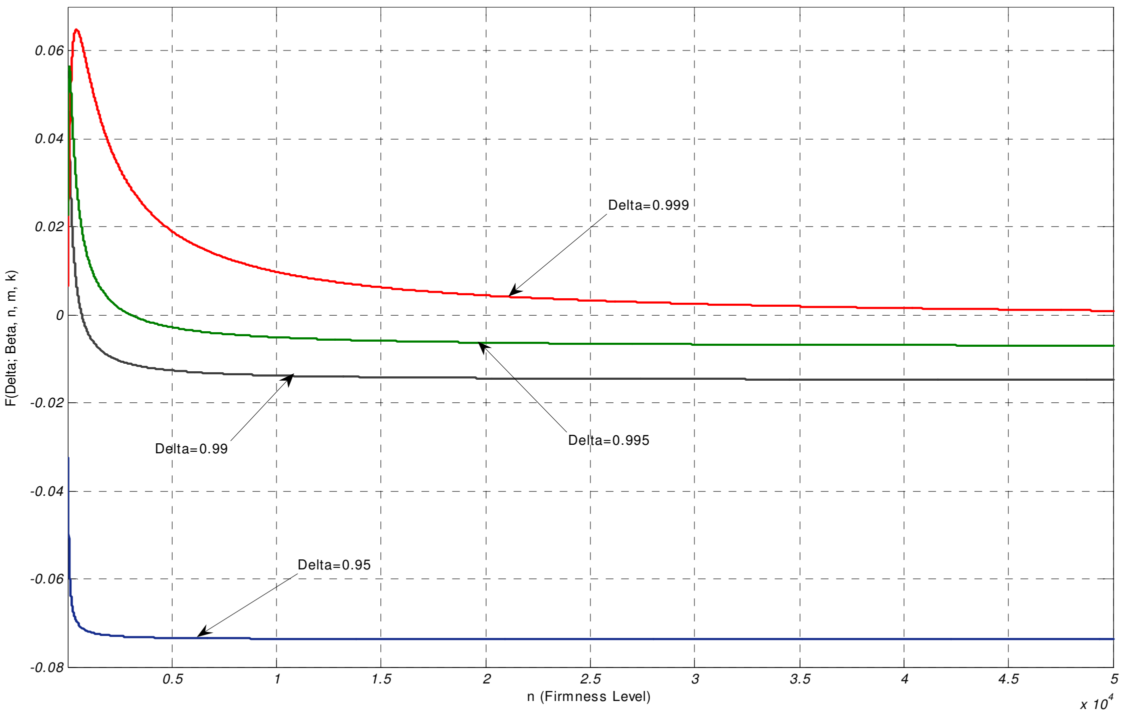
Since n is integer and the right side of *Condition* 1 above may not be integer, we round this value to the smallest integer greater than or equal to it represented by the ceiling function:

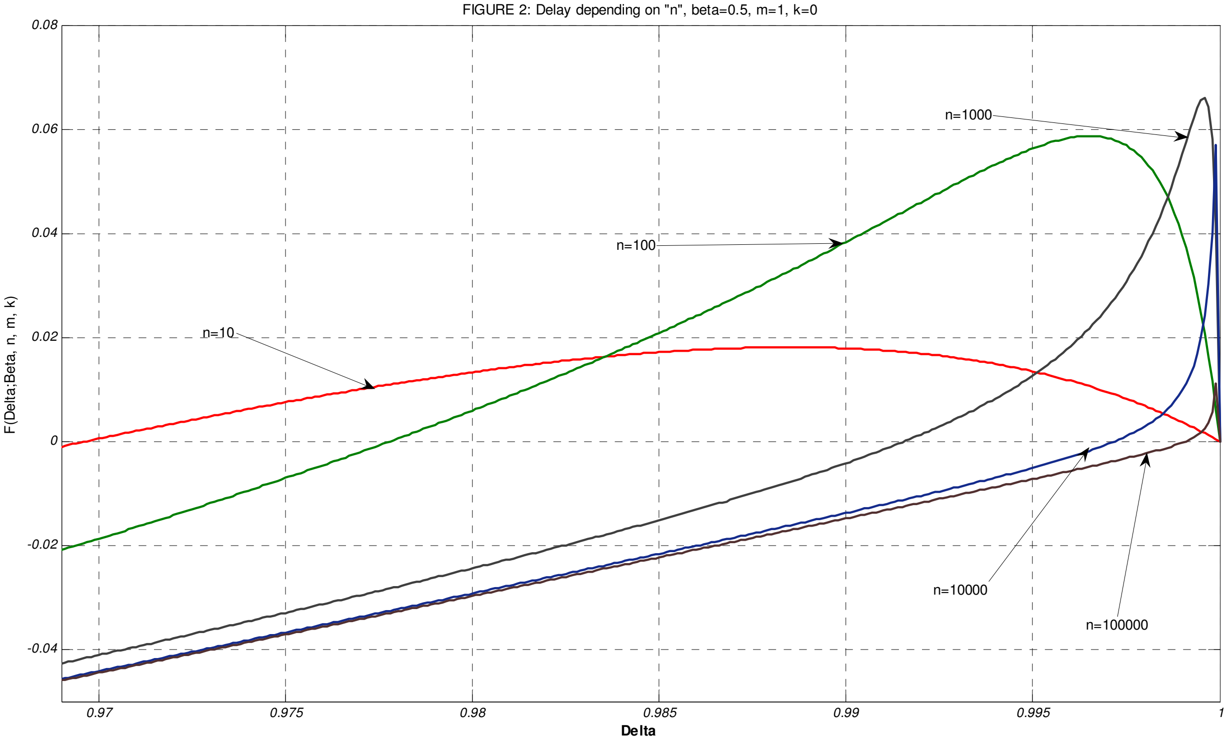
$$n\geq\lceil m_{\beta\delta}+((m_{\beta\delta}+2k)(m_{\beta\delta}+2k+2)(\frac{1-\beta}{\beta}))\rceil$$

Thus, for any given values of $\beta, m_{\beta\delta}$ and n satisfying *Condition* 1, there exists a δ^* such that for any $\delta\geq\delta^*$, we will have $2k+2$ periods delay in the bargaining game between the partially naive agent and the rational agent. ■

Figure 2 shows delay depending on n and some specific parameter values $\beta=0.5$, $m_{\beta\delta}=1$ and $k=0$. As this corollary suggests, whatever the value of n, we can find a patience level δ that will cause delay ($F(\delta; \beta, n, m_{\beta\delta}, k) > 0$, e.g., given $\beta = 0.5$, $m_{\beta\delta} = 1$ and $n = 100$, for all values of δ satisfying $\delta \geq \delta^* = 0.9778$, we get at least 2 period delay- $k = 0$).

FIGURE 1: Delay depending on delta, Beta=0.5, m=1, k=0





6. Conclusion

To be added...

REFERENCES

1. Akin Z., "The Role of Time-Inconsistent Preferences in Intertemporal Investment Decisions and Bargaining", April 2004.

2. Coles Melvyn G. and Muthoo A., Journal of Economic Theory, "Bargaining in a non-stationary environment", March 2003, Volume 109, Issue 1, Pages 70-89.

3. Dellavigna S. and Malmendier U., “Self-Control in the market: Evidence from the Health Club Industry”, mimeo, 2003.

4. Laibson David I., QJE, "Golden eggs and Hyperbolic Discounting.", 1997a, 62: 2, Pages 443-478.

5. O'Donoghue T., Rabin M., AER, "Doing It Now or Later", March 1999a, 89:1, Pages 103-124.

6. O'Donoghue T., Rabin M., QJE, "Incentives for Procrastinators", August 1999b, Pages 769-816.

7. O'Donoghue T., Rabin M., QJE, "Choice and Procrastination", February 2001, 116:1, Pages 121-160.

8. Phelps, E. S. and R. A. Pollak, ReStud, "On Second-Best National Saving and Game-Equilibrium Growth", 1968, 35 (2), 185-199.

9. Rubinstein A., Econometrica, "Perfect Equilibrium in a Bargaining Model", Vol. 50, No. 1. (Jan., 1982), Pages 97-110.

10. Sarafidis Yianis, "Intertemporal Price Discrimination with Time Inconsistent Consumers", January 2004.

11. Shaked A., Sutton J., Econometrica,"Involuntary unemployment as a perfect equilibrium in a Bargaining Model", Vol. 52, No.6, (Nov. 1984), Pages 1351-1364.

12. Strotz Robert H., Review of Economic Studies, "Myopia and Inconsistency in Dynamic Utility Maximization.", 1956, 23, Pages 165-180.

13. Yildiz Muhamet, Econometrica, "Bargaining Without a Common Prior-An Immediate Agreement Theorem", May 2003, 793-811.

14. Yildiz Muhamet, QJE, "Waiting to Persuade", February 2004, 223-248.