ESSAYS ON TIME-INCONSISTENCY AND BARGAINING

A Thesis in
Economics
by
Zafer Akin

© 2006 Zafer Akin

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2006
The thesis of Zafer Akin was reviewed and approved* by the following:

Kalyan Chatterjee  
Distinguished Professor of Economics and Management Science  
Thesis Advisor, Chair of Committee

Susanna Esteban  
Professor of Economics

Edward Green  
Professor of Economics

Abdullah Yavas  
Professor of Business Administration

James S. Jordan  
Professor of Economics  
Acting Head of the Department of Economics

*Signatures are on file in the Graduate School.
Abstract

This thesis consists of three essays focusing on time inconsistency in a game theoretic framework.

The first essay studies an equilibrium concept for both one and two player extensive form games. The equilibrium concept, by construction, embraces different bounded rationality approaches in one framework. A special case that examines the games played by potentially time-inconsistent agents is presented and naive backwards induction solution concept is specified.

The second essay focuses on an alternating offers bargaining game between possibly time inconsistent players. The time inconsistency is modeled by quasi hyperbolic discounting and naive backwards induction solution concept is used in order to obtain the results. Both naive agents who remain naive and those who learn about their own preferences are considered. Offers of the players who are naive are never accepted by any type of player in both no learning and gradual learning cases. The game between a naive player who never learn and a time consistent agent ends in immediate agreement if the time consistent agent is the proposer. A one period delay occurs if time consistent agent is the responder. The more naive the player is, the higher share she gets. In addition, two naive agents who never learn disagree perpetually. When naive agents who are able to learn play against exponential or sophisticated agents, there exists a critical date before which there is no agreement. Therefore, the existence of time inconsistent players who can learn their types as they play the game can be a new explanation for delays in bargaining. The relationship among the degree of naivete, impatience level and bargaining delay is also characterized.

The third essay examines the role of time inconsistency in intertemporal in-
vestment decisions. Quasi hyperbolic discounting framework is incorporated into a game including two stages, namely, self investment game and alternating offers bargaining game. We found that exponential and naive agents finish investment stage without any delay whereas sophisticated agents have a periodical investment schedule. In addition, there exists a threshold level of preference for immediate gratification that makes the sophisticated agent finish investment stage without any delay. When a bonus scheme is added, we showed that the amount of bonus should increase in order for naive agent to continue to invest and finish the investment stage. Moreover, the agents with higher self control problems should be given higher bonus in order to induce them to complete the same investment project.
# Table of Contents

List of Figures vii

List of Tables viii

Acknowledgments ix

Chapter 1
AN EQUILIBRIUM CONCEPT FOR EXTENSIVE FORM GAMES PLAYED BY BOUNDEDLY RATIONAL AGENTS 1
1.1 Introduction ............................................. 1
1.2 The Equilibrium Concept ................................. 2
   1.2.1 Games with only one player ......................... 2
   1.2.2 Games with two players ............................ 13
1.3 A Special Case of Time Inconsistent Preferences .......... 22
1.4 Discussion ................................................. 29
1.5 Conclusion ................................................ 30
1.6 References ............................................... 30

Chapter 2
TIME INCONSISTENCY AND LEARNING IN BARGAINING GAMES 32
2.1 Introduction .............................................. 32
2.2 Solution Concept: Naive Backwards Induction ............... 36
2.3 Model .................................................... 42
2.4 Equilibrium Without Learning ............................ 43
2.5 Equilibrium With Learning ............................... 48
   2.5.1 Immediate Learning ................................. 50
2.5.2 Gradual Learning ........................................ 51
2.6 Conclusion ............................................. 56
2.7 References ............................................. 65

Chapter 3
THE ROLE OF TIME INCONSISTENT PREFERENCES IN INTERTEMPORAL INVESTMENT DECISIONS AND BARGAINING 69
3.1 Introduction ............................................. 69
  3.1.1 The Problem ........................................ 69
  3.1.2 Examples ........................................... 70
  3.1.3 Related Literature .................................. 71
  3.1.4 Contributions of This Paper ....................... 73
  3.1.5 Description of the Model and Preliminary Results .... 76
3.2 The Formal Model ..................................... 78
3.3 Characterizing Equilibrium ............................. 82
  3.3.1 Second Stage Game ................................. 82
  3.3.2 First Stage Game .................................. 82
3.4 Equilibrium With The Bonus Motive .................... 96
3.5 Discussion and Conclusion ............................. 103
3.6 References ............................................. 104
# List of Figures

1.1 Addiction example. .............................................. 4
1.2 Two decision nodes. ............................................. 5
1.3 Three decision nodes. .......................................... 7
1.4 A simple game tree to show how function $m$ works. ........... 10
1.5 Mappings. ....................................................... 12
1.6 Each player has only one decision node. ......................... 16
1.7 One player has one decision node, the other has two decision nodes. 18
1.8 Each player has two decision nodes. ............................ 20
1.9 An example for naive backwards induction. ...................... 27

2.1 An example of naive backwards induction with two different types of opponent. ........................................ 40
2.2 Bargaining shares. ............................................... 48
2.3 Delay depending on delta, beta=0.5, m=1, k=0 .................. 68
2.4 Delay depending on n, beta=0.5, m=1, k=0 ..................... 68

3.1 Optimal investment schedule ................................. 106
List of Tables

1.1 Transition table for two decision nodes. .......................... 6
1.2 Transition table for three decision nodes. .......................... 8
Acknowledgments

I am grateful to my advisor Prof. Kalyan Chatterjee for his continuous guidance, support and advice in preparing my thesis. I am indebted to Prof. Edward Green for his guidance and encouragement especially in preparing the first chapter. I thank Prof. Susanna Esteban and Prof. Abdullah Yavas for their continuous help and support at every stage of my work. I would like to thank participants of Wednesday research meetings at Penn State. I also thank Sophie Bade, James Jordan, Emin Karagozoglu, Vijay Krishna, Selcuk Ozyurt, Yianis Sarafidis, Tomas Sjostrom, Muhamet Yildiz, Ruilin Zhou for helpful comments. All errors remain my own.
Dedication

This thesis is dedicated to my beloved family and my friends who made State College a heaven on earth and leaves me unforgettable memories...
1.1 Introduction

Consider an agent who has the choices of going home directly after work or going to a bar with friends and having a drink before going home. He may have different thoughts such as “if I go to a bar and have a drink, I may end up getting drunk which I do not want” or “there is no harm in having just one drink, then I can go home, no problem!” With these thoughts, the agent may end up directly going home or getting drunk or going home after one drink. It really depends on what kind of state of mind the agent has. What should this person do? Given his state of mind at 5.30pm (end of the day at work), he will compare the options he has by taking into account the implications of his state of mind and choose the optimal one. “Nothing happens with a drink,” “I will probably be regretful if I go to a bar, let me go home,” “I will probably be regretful if I go to a bar, but let me go,” “let me go and see what happens,” these are all different behavioral outcomes implied by different states of mind.

There are different models in the fields of Economics that deal with these kind of situations, especially the bounded-rationality models. Regret models, addiction, procedural rationality and present bias with sophisticated and naive beliefs are some examples. All of these models are motivated by limitations that the
economic agents face (cognitive, ability-based etc.) and/or inherited behavioral characteristics of the agents.

In this paper, we focus on a very general approach that takes its underlying motivation from *the state of mind specification*. State of mind is a very general language and a comprehensive enough concept that covers any situation that one can think of (not only the economic decision problems, as the one above) and potentially embraces all the motivating figures of the models mentioned. By taking this specification as a benchmark, we propose a new equilibrium concept for a restricted class of one and two player extensive form games. As a first step, we take a simple and a deterministic approach that has its own dynamics and that can easily be extended to more general frameworks. A formalization of the motivating example and what we mean by more general frameworks will be discussed shortly.

The rest of the paper is organized as follows. Section 2 discusses the equilibrium concept in detail for a restricted class of one player and two player games in order. Section 3 introduces a special case where the games played by potentially time-inconsistent agents are examined. Section 4 discusses some features of the existing model and an alternative approach as an extension. Section 5 concludes.

### 1.2 The Equilibrium Concept

#### 1.2.1 Games with only one player

Let $\Gamma$ be a generic extensive form game. There is only one player. A game tree is defined by the following objects: $(N, A, Z)$ where $N$ : \{nodes\}, $A$ : \{actions\}, $Z$ : \{terminal nodes\}. Let $X$ be the set of decision nodes and $x_0 \in X$ be the initial node, $X = N \setminus Z$, $N = X \cup Z$. We assume perfect information (each information set is singleton). Let $a(x)$ be the set of actions from $x$ for $x \in X$. Each node is reached by a single path. Let $\Gamma^x$ be the subgame whose initial node is $x$, $x \in X$. Define $N^x$ as the set of nodes including $x$ and its successors. $X^x$ and $Z^x$ are defined by restrictions of $X$ and $Z$ to $N^x$. Define $A = \bigcup_{x \in X} a(x)$. Let $P(x)$ represent the set of predecessors of node $x$ and $S(x) = X^x \setminus x$ is the set of all successors of node $x$.

Let $\Pi$ be the set of agents’ possible states of mind and it is inclusive enough that it captures all the different types of beliefs of the agents. Let $\pi_{x_0} \in \Pi$ be the
initial state of mind (state of mind at node \(x_0\)) and it is given. Let \(u\) be the utility function and \(u : \Pi \times Z \rightarrow \mathbb{R}\). Let \(\Psi\) be the conjectured transition function of the states of minds and \(\Psi : \Pi \times A \rightarrow \Pi\). This means that the agent has potentially a different state of mind at each decision node \(x \in X\) whether he moves at that node or not. Finally, let \(\overline{\Psi}\) be the true transition function of the agents.

Some specifications:

a. Initial state of the agent, \(\pi_{x_0} \in \Pi\), is given such that \(u(\pi_{x_0}, z), \forall z \in Z\) is the utility function at \(x_0\) and \(\Psi(\pi_{x_0}, a), \forall a \in a(x_0)\) is the conjectured transition function.

b. At any node \(\hat{x}\), let \((x', x'') = a\) be an action in \(a(x')\) where \(x', x'' \in S(\hat{x})\), then

The agent is in state \(\pi_{x'}\) and \(\pi_{x''} = \Psi(\pi_{x'}, a)\) where \(\pi_{x'}\) and \(\pi_{x''}\) are the conjectured states at node \(x'\) and \(x''\), respectively, from the perspective of node \(\hat{x}\) and they are determined by \(\Psi(., .)\) recursively. Moreover, the agent at \(\hat{x}\) thinks that \(\pi_{x''}\) will be the state at node \(x''\) from the perspective of all the nodes \(x \in X(\hat{x}) \cap P(x'')\). In addition, at the node \(x'\), the actual state will be \(\pi_{x'}(i)\) that is determined by \(\overline{\Psi}(., .)\) recursively and \(\pi_{x''} = \Psi(\pi_{x'}, a)\).

c. \(\pi_{x''} = \overline{\Psi}(\pi_{x'}, a)\)

This specification means that at any node, the agent realizes his own true state and his true state in the next node is given by the true transition function, \(\overline{\Psi}(., .)\).

Current and future states are determined based on the above specifications. At each of his decision nodes \(x\), given the state, the conjectured utility function and the transition function, the agent predicts how the game will proceed. Depending on this, he determines the conjectured path leading to the terminal node \((z^x)^*\). Then, at the node \(x\), he chooses the action \(a^* \in a(x)\) associated with this path.

Now, we turn to the example given at the beginning. In Figure 1.1, assume that there are three states, \(\pi_{01}\), \(\pi_{02}\) and \(\pi_1\). Conjectured transitions and actual transitions are given as follows:

\[
\Psi(\pi_{01}, a) = \pi_{01}; \quad \Psi(\pi_{02}, a) = \pi_1; \quad \overline{\Psi}(\pi_{01}, a) = \overline{\Psi}(\pi_{02}, a) = \pi_1
\]
Utilities depending on the states satisfy the following:

\[ u(\pi_{01}, z_3) > u(\pi_{01}, z_1) > u(\pi_{01}, z_2) \]
\[ u(\pi_{02}, z_3) > u(\pi_{02}, z_1) > u(\pi_{02}, z_2) \]
\[ u(\pi_1, z_1) > u(\pi_1, z_2) > u(\pi_1, z_3) \]

This example illustrates two addicts 1 and 2 having states \( \pi_{01} \) and \( \pi_{02} \), respectively. Initially, both prefer having a drink before going home, \( z_3 \), to directly going home, \( z_1 \), that is preferred to get drunk, \( z_2 \). Agent 1 thinks that his preference ordering will not change if he goes to a bar, \( \Psi(\pi_{01}, a) = \pi_{01} \). Agent 2 thinks that his preference ordering will change if he goes to a bar and he will keep drinking, \( \Psi(\pi_{02}, a) = \pi_1 \). So, for the second agent, \( (b, c) \) strategy is optimal that is going home directly. On the other hand, the first agent will go to a bar with the hope of having only one drink and then going home but he will end up getting drunk. It seems \( (a, d) \) strategy is optimal for him but he ends up following \( (a, c) \) which is the worse outcome according to his original state of mind\(^1\).

\(^1\)Incorporating regret motives and probabilistic transitions would be a plausible extension.
Note that in the above specification, the states are not themselves the discounted utility or the beliefs about the future preferences. Instead, these are indicated by different states in the state (type) space $\Pi$.

Note also that at a given node, the agent has a given state. If it is the initial node, the state is exogenously given; if it is a decision node, the state is recursively assigned to that node by the conjectured transition function. This state implies what his utility function will be at that node. When he looks ahead and determines the future states from the perspective of any given node $x$, his perception on what his future states will be at each successor of $x$ is based on the transition function and the state he has at node $x$.

A Restricted Class of Perfect Information Extensive-Form Games

**STEP 1: Games with 2 decision nodes**

![Figure 1.2. Two decision nodes.](image)

In Figure 1.2, at $x_0$, agent believes that action $a$ will result in his having state of mind $\Psi(\pi_{x_0}, a)$ at $x_1$ and action $b$ will result in his having state of mind $\Psi(\pi_{x_0}, b)$ at $z_1$ and so that his preferences at $x_1$ and $z_1$ would be $u(\Psi(\pi_{x_0}, a), .)$ and $u(\Psi(\pi_{x_0}, b), .)$, respectively. However, when he reaches to these nodes, he will
actually have preferences $u(\Psi(\pi_{x_0}, a), .)$ and $u(\Psi(\pi_{x_0}, b), .)$, respectively.

Assume that the agent does not gain any utility unless the game ends (there is no associated utility with any of the decision nodes). Let $w_i$ be the outcome he gets if terminal node $z_i$ is reached. Note that distinct nodes may have the same outcome. A special case is worth mentioning here. This is an agent where he has $(\delta, \beta, \hat{\beta})$ where $\delta$ is time-consistent discount factor, $\beta$ is time-inconsistent impatience and $\hat{\beta}$ is his perception about the value of his $\beta$ in the future. The intertemporal preferences for this special case can be represented by the following utility function:

At $x_0$, he has: $u(\pi_{x_0}, z_1) = \beta \delta w_1$; $u(\pi_{x_0}, z_2) = \beta \delta^2 w_2$; $u(\pi_{x_0}, z_3) = \beta \delta^2 w_3$

At $x_0$, he thinks at $x_1$ he will have: $u(\Psi(\pi_{x_0}, a), z_2) = \hat{\beta} \delta w_2$; $u(\Psi(\pi_{x_0}, a), z_3) = \hat{\beta} \delta w_3$

At $x_1$, he will actually have: $u(\Psi(\pi_{x_0}, a), z_2) = \beta \delta w_2$; $u(\Psi(\pi_{x_0}, a), z_3) = \beta \delta w_3$

This case will be examined later in detail. However, note that our definition allows more general cases.

Actual and conjectured transitions are shown in the following table:

<table>
<thead>
<tr>
<th></th>
<th>$x_0$</th>
<th>$x_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_0$</td>
<td>$\pi_{x_0}$</td>
<td>$\Psi(\pi_{x_0}, a)$</td>
</tr>
<tr>
<td>$x_1$</td>
<td>$\pi_{x_0}$</td>
<td>$\Psi(\pi_{x_0}, a)$</td>
</tr>
</tbody>
</table>

Table 1.1 shows the conjectured state of the agent at the column nodes from the perspective of the row nodes. For example, for the $x_0 - x_1$ entry in the table, the player is at node $x_0$ and he conjectures that his state at $x_1$ will be the corresponding entry, $\Psi(\pi_{x_0}, a)^2$. Together with the above specifications, we can make the following definition of this game’s equilibrium:

Definition 1: A noncooperative equilibrium of the game above, defined by the set of states, $\Pi$, the set of actions, $A$, the utility function, $u$, the conjectured

---

Note that lower diagonal elements are same with the diagonal element of the column that they belong to. In other words, $a_{ij} = a_{jj}$ $\forall i > j$ where $a_{ij}$ represents the element located at the row $i$ and column $j$. This means that the agent remembers what his state was at all the previous nodes.
transition function, $\Psi$, and the true transition function of states, $\overline{\Psi}$, is a path from the initial node to a terminal node such that the initial state is $\pi_{x_0}$ at the initial node $x_0$ and the terminal node is given by

$$(z^{x_0})^* = \arg \max \{ u(\pi_{x_0}, z_1), u(\pi_{x_0}, \arg \max \{ u(\Psi(\pi_{x_0}, a), z_2), u(\Psi(\pi_{x_0}, a), z_3) \}) \}$$

The actual path in the game does not have to be the same with the initially perceived one. The actual path will be the collection of actions such that

* Actual and conjectured state at node $x_1$ is determined based on the above specification.
* At each decision node $x \in X$, given the state and the conjectured transition function, the agent chooses the action $a^* \in a(x)$ associated with $(z^x)^*$ where

$$(z^{x_0})^* = \arg \max \{ u(\pi_{x_0}, z_1), u(\pi_{x_0}, \arg \max \{ u(\Psi(\pi_{x_0}, a), z_2), u(\Psi(\pi_{x_0}, a), z_3) \}) \}$$

If $(z^{x_0})^* \neq z_1$ or $x_1$ is reached, then

$$(z^{x_1})^* = \arg \max \{ u(\overline{\Psi}(\pi_{x_0}, a), z_2), u(\overline{\Psi}(\pi_{x_0}, a), z_3) \}$$

**STEP 2: Games with three decision nodes**

![Three decision nodes](image)

**Figure 1.3.** Three decision nodes.
In Figure 1.3, at $x_0$, agent believes that taking actions $a$ and $b$ will result in his having state of mind $\Psi(\pi_{x_0}, a)$ and $\Psi(\pi_{x_0}, b)$ at $x_1$ and $z_1$, so that his preferences at $x_1$ and $z_1$ would be $u(\Psi(\pi_{x_0}, a), .)$ and $u(\Psi(\pi_{x_0}, b), .)$, respectively. However, when he reaches to these nodes, he will actually have preferences $u(\overline{\Psi}(\pi_{x_0}, a), .)$ and $u(\overline{\Psi}(\pi_{x_0}, b), .)$, respectively. He also believes that taking actions $c$ and $d$ will result in his having state of mind $\Psi(\Psi(\pi_{x_0}, a), c)$ and $\Psi(\Psi(\pi_{x_0}, a), d)$ at $x_2$ and $z_2$, respectively.

Again assume that the agent does not gain any utility unless the game ends (there is no associated utility with any of the decision nodes). Let $w_i$ be the outcome he gets if terminal node $z_i$ is reached. Actual and conjectured states are shown in Table 1.2.

**Table 1.2.** Transition table for three decision nodes.

<table>
<thead>
<tr>
<th></th>
<th>$x_0$</th>
<th>$x_1$</th>
<th>$x_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_0$</td>
<td>$\pi_{x_0}$</td>
<td>$\Psi(\pi_{x_0}, a)$</td>
<td>$\Psi(\Psi(\pi_{x_0}, a), c)$</td>
</tr>
<tr>
<td>$x_1$</td>
<td>$\pi_{x_0}$</td>
<td>$\Psi(\pi_{x_0}, a)$</td>
<td>$\Psi(\Psi(\pi_{x_0}, a), c)$</td>
</tr>
<tr>
<td>$x_2$</td>
<td>$\pi_{x_0}$</td>
<td>$\Psi(\pi_{x_0}, a)$</td>
<td>$\Psi(\Psi(\pi_{x_0}, a), c)$</td>
</tr>
</tbody>
</table>

Together with the above specifications, we can make the following definition of this game’s equilibrium:

**Definition 2:** A noncooperative equilibrium of the three decision node game, defined by the set of states, $\Pi$, the set of actions, $A$, the utility function, $u$, the conjectured transition function, $\Psi$, and the true transition function of states, $\overline{\Psi}$, is a path from the initial node to a terminal node such that the initial state is $\pi_{x_0}$ at the initial node $x_0$ and the terminal node is given by

$$(z^{x_0})^* = \arg\max\{u(\pi_{x_0}, z_1), u(\pi_{x_0}, \arg\max\{u(\Psi(\pi_{x_0}, a), z_2), u(\Psi(\pi_{x_0}, a), (z^{x_2})')\})\}$$

where $(z^{x_2})' = \arg\max\{u(\Psi(\pi_{x_0}, a), c), z_3), u(\Psi(\pi_{x_0}, a), c), z_4)\}$

The actual path in the game does not have to be the same with the initially perceived one. The actual path will be the collection of actions such that

* Actual and conjectured state at node $x_1$ and $x_2$ is determined based on the above specification.
At each decision node \( x \in X \), given the state and the conjectured transition function, the agent chooses the action \( a^* \in a(x) \) associated with \((z^*)^*\) where

\[
(z^x)^* = \arg \max \{ u(\pi_{x0}, z_1), u(\pi_{x0}, \arg \max \left\{ u(\Psi(\pi_{x0}, a), z_2), u(\Psi(\pi_{x0}, a), (z^{x2})') \right\}) \}
\]

where \((z^{x2})' = \arg \max \{ u(\Psi(\Psi(\pi_{x0}, a), c), z_3), u(\Psi(\Psi(\pi_{x0}, a), c), z_4) \}\}

If \((z^x)^* \neq z_1 \) or \( x_1 \) is reached, then

\[
(z^x_1)^* = \arg \max \{ u(\Psi(\pi_{x0}, a), z_2), u(\Psi(\pi_{x0}, a), \arg \max \left\{ u(\Psi(\pi_{x0}, a), c), z_3 \right\}), u(\Psi(\Psi(\pi_{x0}, a), c), z_4) \} \}
\]

If \((z^x_1)^* \neq z_2 \) or \( x_2 \) is reached, then

\[
(z^x_2)^* = \arg \max \{ u(\Psi(\Psi(\pi_{x0}, a), c), z_3), u(\Psi(\Psi(\pi_{x0}, a), c), z_4) \}
\]

**STEP T: Games with ”T” decision nodes**

Let \( \Gamma \) be a \( T \) period extensive form game. There is only one player. To define a noncooperative equilibrium of \( \Gamma \), we define two mappings \( m \) and \( g \).

Function \( m \), given any node, assigns a state to all successors of that node for all the nodes in the game, \( m : X \times N \rightarrow \Pi \) and is formally defined as

\[
m(x, y) = \begin{cases} 
\pi_x, & \text{if } y = x \\
\Psi(m(x, p(y)), a(p(y), y)), & \text{if } y \in N^x \setminus \{x\}
\end{cases}
\]

where \( a(x, y) \) is the action from node \( x \) that leads to node \( y \); \( p(y) \) is the immediate predecessor of \( y \). In other words, \( m(x, y) \) assigns a state to every node, \( y \in N^x \), in the subgame starting with the node \( x \) from the perspective of node \( x \). Note that if \( y = x \), then the state that he assigns to the node where he is at is the true state, \( \pi_x \), of himself. Note also that if \( y \) is an immediate successor of \( x \), then \( p(y) = x \) and

\[
m(x, y) = \Psi(m(x, p(y)), a(p(y), y)) = \Psi(m(x, x), a(x, y)) = \Psi(\pi_x, a(x, y)).
\]

Note that the true state at any given node is exogenously given if it is the initial node, if not, it is determined by the true transition function \( \Psi(., .) \). In the definition of \( m(x, y) \), we did not specify what the agent’s state is at node \( x \). If the agent is actually at node \( x \) and assigning a state to a node \( y \in N^x \), then his actual
state at \( x \) is given by \( m(x_0, x) = \Psi(\Psi(\Psi(...\Psi(\pi_{x_0},..)),..),..) \). This is different than the case where the agent is actually at some node \( x' \in P(x) \) and assigning a state to node \( x \) (If \( x' = x_0 \), \( \pi_{x'} = \pi_{x_0} \) is given; If \( x' \neq x_0 \), \( \pi_{x'} \) is determined by \( \Psi \) recursively.) In that case, the conjectured state at \( x \) will be given by \( m(x', x) = \Psi(\Psi(...\Psi(\Psi(\pi_{x'},..),..),..),..) \).

To see how function \( m \) works, we will give the following example. We want to find, at node \( x \), what the conjectured state about node \( y_3 \) is (see Figure 1.4). This is given by \( m(x, y_3) \).

![Figure 1.4. A simple game tree to show how function \( m \) works.](image-url)

From the definition, we can write it as

\[
m(x, y_3) = \Psi(m(x, y_2), a(y_2, y_3))
\]

Now, we should find \( m(x, y_2) \);

\[
m(x, y_2) = \Psi(m(x, y_1), a(y_1, y_2))
\]
To find \( m(x, y_2) \), we need to find \( m(x, y_1) \);

\[
m(x, y_1) = \Psi(\pi_x, a(x, y_1))
\]

Now, plug these recursively into \( m(x, y_3) \) to get:

\[
m(x, y_3) = \Psi(\Psi(\Psi(\pi_x, a(x, y_1)), a(y_1, y_2)), a(y_2, y_3))
\]

Function \( g : \Pi \times N \rightarrow \mathbb{Z} \) assigns a terminal node to all the nodes in the game given the state of the agent at that node. Suppose \( \forall \pi \in \Pi, \forall y \in s(x), g(\pi, y) \) is defined where \( s(x) \) is the set of immediate successors of node \( x \). Then, \( g \) is defined as follows:

\[
g(\pi_x, x) = \{ \arg \max_{y \in s(x)} \{ u(\pi_x, g(m(x, y), y)) \} , \text{ if } x \in \mathbb{Z}
\]

Simply, function \( g(\cdot, \cdot) \) assigns a terminal node to every decision node. If \( \pi_x = \pi_x \) ( \( \pi_x \) : the actual state at node \( x \) implied by the true transition function, \( \Psi \) ), then call \( g(\pi_x, x) \) as \( g_x \). If \( \pi_x = m(x', x) \) ( \( m(x', x) \) : the conjectured state at node \( x \) from the perspective of node \( x' \in P(x) \), implied by the conjectured transition function, \( \Psi \) ), then call \( g(\pi_x, x) \) as \( g_x \). Let \( a_{x_i}(x_j) \) be the action in \( a(x_j) \) associated with the terminal node assigned to \( x_j \) by \( g(\pi_x, x_j) \) from the perspective of node \( x_i \) where \( \pi_{x_j} = m(x_i, x_j) \).

Based on the above definitions, a noncooperative equilibrium of a one player game can be defined as follows:

**Definition 3:** A noncooperative equilibrium of a one player (finite period, extensive form) game, defined by the set of states, \( \Pi \), the set of actions, \( A \), the utility function, \( u \), the transition function, \( \Psi \), and the true transition function of states, \( \Psi \), is a path from the initial node to a terminal node such that the initial state is \( \pi_{x_0} \) at the initial node \( x_0 \) and the terminal node is \( g_{x_0} = g(\pi_{x_0}, x_0) \) defined above. Again the conjectured path from \( x_0 \) to \( g_x \) may not overlap with the actual path that will arise during the course of the game. This is true for any decision node \( x \in X \). When the game proceeds, the agent’s state of mind evolves based on
the true transition function, so the actions that are associated with the terminal nodes $\overline{g}_{x_1}$, $\overline{g}_{x_2}$... from the nodes that are reached will be chosen. In other words, at any node $x$, the agent chooses the action $a_x(x) \in a(x)$ associated with $\overline{g}_x$ and $(x, a_x(x)) = y \in s(x)$ and $\overline{g}_x = g(\pi_x, x) = g(\pi_y, y) = g(m(x, y), y) = z^* \in \mathbb{Z}^x$ where $y$ is an element of the set of decision nodes on the conjectured equilibrium path. However, at node $y$, it may be the case that $\overline{g}_y = g(\pi_y, y) = z^{**} \neq z^* \in \mathbb{Z}^x$.

Note that function $m$ assigns a state to each node $x$ by taking the transition function into account. The assignment of a state to node $x$ by function $m$ implicitly means that the utility function assigns a real number to the terminal nodes from the perspective of node $x$. Then, function $g$ assigns a terminal node $z^*$ to node $x$. The mappings are summarized in Figure 1.5.

![Figure 1.5. Mappings.](image)

To see how the function $g$ recursively assigns every node to a terminal node, we will examine $T = 3$ case. The equilibrium for this game is the path from $x_0$ to $\overline{g}_{x_0} = g(\pi_{x_0}, x_0)$. We now find

$$
\overline{g}_{x_0} = \arg \max_{y \in s(x_0)} \{ u(\pi_{x_0}, g(m(x_0, y), y)) \}$$
\( g_{x_0} = \arg \max \{ u(\pi_{x_0}, z_1), u(\pi_{x_0}, g(m(x_0, x_1), x_1)) \} \)

To find the \( g_{x_0} \), we need to find first \( g_{x_1} = g(m(x_0, x_1), x_1) \) as follows:

\[ g_{x_1} = \arg \max_{y \in s(x_1)} \{ u(m(x_0, x_1), g(m(x_0, y), y)) \} \]

\[
\begin{align*}
g_{x_1} &= \arg \max \{ u(\Psi(\pi_{x_0}, a_{x_1}(x_0)), z_2), u(\Psi(\pi_{x_0}, a_{x_1}(x_0)), g(m(x_0, x_2), x_2)) \}
\end{align*}
\] (2)

To find the \( g_{x_1} \), we need to go one step further and find \( g_{x_2} = g(m(x_0, x_2), x_2) \) as follows:

\[ g_{x_2} = \arg \max_{y \in s(x_2)} \{ u(\Psi(\pi_{x_0}, a_{x_1}(x_0)), a_{x_2}(x_1)), g(m(x_0, y), y)) \} \]

\[ g_{x_2} = \arg \max \{ u(\Psi(\pi_{x_0}, a_{x_1}(x_0)), a_{x_2}(x_1)), z_3), u(\Psi(\pi_{x_0}, a_{x_1}(x_0)), a_{x_2}(x_1)), z_4) \} \]

Then, plug \( g_{x_2} \) in \( g_{x_1} \) and get

\[ g_{x_1} = \arg \max \{ u(\Psi(\pi_{x_0}, a_{x_1}(x_0)), z_2), u(\Psi(\pi_{x_0}, a_{x_1}(x_0)), \arg \max \{ u(\Psi(\pi_{x_0}, a_{x_1}(x_0)), a_{x_2}(x_1)), z_3), u(\Psi(\pi_{x_0}, a_{x_1}(x_0)), a_{x_2}(x_1)), z_4) \} \}

Then, plug \( g_{x_1} \) in \( g_{x_0} \) and get \( g_{x_0} = g(\pi_{x_0}, x_0) \)

\[ g_{x_0} = g(\pi_{x_0}, x_0) = \arg \max \{ u(\pi_{x_0}, z_1), u(\psi_{x_0}, \arg \max \{ u(\Psi(\pi_{x_0}, a_{x_1}(x_0)), z_2), u(\Psi(\pi_{x_0}, a_{x_1}(x_0)), \arg \max \{ u(\Psi(\pi_{x_0}, a_{x_1}(x_0)), a_{x_2}(x_1)), z_3), u(\Psi(\pi_{x_0}, a_{x_1}(x_0)), a_{x_2}(x_1)), z_4) \} \}

Thus, the mapping \( g \) allows us to extend the definition of the equilibrium to the games with \( T \) decision nodes.

### 1.2.2 Games with two players

Let \( \Gamma \) be a generic extensive form game. There are two players, \( i \neq j \in \{1, 2\} \). A game tree is defined by the following objects: \( (N, A, Z) \) where \( N : \{\text{nodes}\}, A : \{\text{actions}\}, Z : \{\text{terminal nodes}\} \). Let \( X \) be the set of decision nodes, \( x_0 \in X \) be
the initial node and $X = N \setminus Z$. $N = X \cup Z$. We assume perfect information (each information set is singleton). The function $\lambda$ assigns each nonterminal node to a player, $\lambda : X \rightarrow \{i, j\}$. Let $a(x)$ be the set of actions from $x$ for $x \in X$. In other words, $a(x)$ is the set of moves for $\lambda(x)$ at $x$. Each node is reached by a single path. Let $\Gamma^x$ be the subgame whose initial node is $x$ for $x \in X$. Define $N^x$ as the set of nodes including $x$ and its successors. $X^x$ and $Z^x$ are defined by restrictions of $X$ and $Z$ to $N^x$. Define $A = \bigcup_{x \in X} a(x)$. Let $P(x)$ represent the set of all predecessors of node $x$ and $S(x) = X^x \setminus x$ is the set of all successors of node $x$.

Let $\Pi$ be the set of agents’ possible states of mind and it is inclusive enough that it captures all the different types of beliefs of the agents. Let $\pi_{\pi^i_{x_0}(i)} \in \Pi$ be the initial state of mind of agent $i$ and it is privately known by agent $i$. Let $u$ be the utility function and $u : \Pi \times Z \rightarrow \mathbb{R}$. Let $\Psi$ be the conjectured transition function of the states of minds and $\Psi : \Pi \times A \rightarrow \Pi$. This means that each agent has potentially a different state of mind at each decision node $x \in X$ whether he moves at that node or not. Let $l$ be a function that maps state space to itself, $l : \Pi \rightarrow \Pi$. In words, $l(\pi(i)) = \pi(j)$ is $i$’s conjecture about what $j$’s state is.

Finally, let $\bar{\Psi}$ be the true transition function of the agents. Function $u$, $l$ and $\Psi$ are common knowledge but $\bar{\Psi}$ is not.

Some specifications:

a. Initial state of the agent $i$, $\pi_{\pi^i_{x_0}(i)} \in \Pi$, is given such that $u(\pi_{x_0}(i), z)$, $\forall z \in Z$ is the utility function of agent $i$ at $x_0$; $\Psi(\pi_{x_0}(i), a)$, $\forall a \in a(x_0)$ is the conjectured transition function of agent $i$ and $l(\pi_{x_0}(i))$ is the function that assigns a state to agent $j$ given $i$’s initial state $\pi_{x_0}(i)$.

b. $l(\pi_{x_0}(i)) = \pi^i_{x_0}(j)$: it’s conjecture about the initial state of $j$,

$$u(\pi^i_{x_0}(j),.) : it’s conjecture about the valuation of terminal nodes of agent $j$ from the perspective of node $x_0$,

$$\Psi(\pi^i_{x_0}(j),.) : it’s conjecture about agent $j$’s states at the immediate successors of $x_0$ from the perspective of node $x_0$.

c. At any node $\hat{x}$, let $(x', x'') = a$ be an action in $a(x')$ where $x', x'' \in X^{\hat{x}}$, then $\Psi(\pi_{x'}(i), a)$ where $\pi_{x'}(i)$ and $\pi_{x''}(i)$ are the conjectured states at node $x'$ and $x''$ agent $i$, respectively, from the perspective of node $\hat{x}$ and they are determined by $\Psi(.,.)$ recursively. Moreover, the agent $i$ at $\hat{x}$ thinks that $\pi_{x''}(i)$ will be his state at node $x''$ from the perspective of all the
nodes $x \in X^2 \cap P(x'')$. In addition, at the node $x'$, the actual state of the agent $i$ will be $\pi_{x'}(i)$ that is determined by $\Psi(\cdot, \cdot)$ recursively and $\pi_{x''}(i) = \Psi(\pi_{x'}(i), a)$.

d. At any node $\hat{x}$, let $(x', x'') = a$ be an action in $a(x')$ where $x', x'' \in X^2$, then Agent $i$ is in state $\pi_{\hat{x}}(i)$. This implies $l(\pi_{\hat{x}}(i)) = \pi_{\hat{x}}^i(j)$. Furthermore, $\pi_{x''}^i(j) = \Psi(\pi_{x'}^i(j), a)$ where $\pi_{x'}^i(j)$ is $i$’s conjecture about the state of agent $j$ at node $x'$ from the perspective of node $\hat{x}$ and it is determined by $\Psi(\cdot, \cdot)$ recursively and $\pi_{x''}^i(j)$ is the node $\hat{x}$ conjecture of the agent $i$ about node $x''$ state of agent $j$. Moreover, agent $i$ at $\hat{x}$ thinks that $\pi_{x''}^i(j)$ will be the $j$’s state at node $x''$ from the perspective of all the nodes $x \in X^2 \cap P(x'')$. In addition, at the node $x'$, the actual state of the agent $i$ will be $\pi_{x'}(i)$ that is determined by $\Psi(\cdot, \cdot)$ recursively and then $l(\pi_{x'}(i)) = \pi_{x'}(j)$ and $\pi_{x''}^i(j) = \Psi(\pi_{x'}^i(j), a)$. Moreover, both $\pi_{x'}^i(j)$ and $\pi_{x''}^i(j)$ are potentially different than the ones conjectured at node $\hat{x}$.

e. $\pi_{x''}(i) = \Psi(\pi_{x'}(i), a)$

This specification means that at any node, each agent realizes his own true state and his true state in the next node is given by the true transition function, $\Psi(\cdot, \cdot)$.

True and conjectured states are determined based on the above specifications. Each agent, at each of the decision node $x$, given his state, his conjectured utility function, transition function and the conjectured state of the opponent, predicts how the game will proceed. Depending on this, he determines the conjectured path leading to the terminal node $(z^x)^*$. Then, at the node $x$, he chooses the action $a^* \in a(x)$ associated with this path.

Again note that when the agent $i$ looks ahead from a node $x$ and determines his own future states, his perception on what his future states will be at each successor of $x$ is based on the transition function and the state he has at node $x$, $\Psi(\pi_x(i), \cdot)$. His perception on what his opponent’s state will be at each successor of $x$ is based on the transition function and the state that he assigns to his opponent at node $x$, $l(\pi_x(i)) = \pi_x^i(j)$ and $\Psi(\pi_x^i(j), \cdot)$.

Without loss of generality, in every case, we will assume that player 1 is the first mover.

A Restricted Class of Perfect Information Extensive-Form Games

**STEP 1:** Each player has only one decision node
Initial states of agent 1 and 2 are $\pi_{x_0}(1)$ and $\pi_{x_0}(2)$, respectively. This implies that taking action $a$ ($b$) will cause agent 1 to have state of mind $\Psi(\pi_{x_0}(1), a)$ at $x_1$ ($\Psi(\pi_{x_0}(1), b)$ at $z_1$). The initial state of 1, $\pi_{x_0}(1)$, also implies that according to agent 1, agent 2 has state $l(\pi_{x_0}(1)) = \pi_{x_0}^1(2)$ and has utility function $u(\pi_{x_0}^1(2), \cdot)$ at node $x_0$.

Assume that the agent does not gain any utility unless the game ends (there is no associated utility with any of the decision nodes). Let $w_k = (w_{k1}, w_{k2})$, where $w_{ki}$ is the outcome of agent $i$, be the outcome vector if terminal node $z_k$ is reached. Note that distinct nodes may have the same outcome. Together with the above specifications, we can make the following definition of this game’s equilibrium:

**Definition 4**: A noncooperative equilibrium of the game with each player having one decision node, defined by the set of states, $\Pi$, the set of actions, $A$, state dependent utility function, $u(\cdot, \cdot)$, the transition function, $\Psi(\cdot, \cdot)$, the function $l(\cdot)$ and the true transition function, $\overrightarrow{\Psi}(\cdot, \cdot)$, is a path from the initial node to a terminal node such that the initial state is $\pi_{x_0}(i)$ at the initial node $x_0$ and the terminal
node is given by

\[(z^{x_0})^* = \arg \max \{ u(\pi_{x_0}(1), z_1), u(\pi_{x_0}(1), \arg \max \{ u(\Psi(l(\pi_{x_0}(1)), a), z_2), u(\Psi(l(\pi_{x_0}(1)), a), z_3) \}) \} \]

The actual path in the game does not have to be the same with the initially perceived one. The actual path will be the collection of actions such that

* Actual and conjectured state at node \( x_1 \) is determined based on the above specification.
* At each decision node \( x \in X \), agents choose the action \( a^* \in a(x) \) associated with \((z^x)^*\) where

\[(z^{x_0})^* = \arg \max \{ u(\pi_{x_0}(1), z_1), u(\pi_{x_0}(1), \arg \max \{ u(\Psi(l(\pi_{x_0}(1)), a), z_2), u(\Psi(l(\pi_{x_0}(1)), a), z_3) \}) \} \]

If \((z^{x_0})^* \neq z_1\) or \( x_1 \) is reached, then

\[(z^{x_1})^* = \arg \max \{ u(\overline{\Psi}(\pi_{x_0}(2), a), z_2), u(\overline{\Psi}(\pi_{x_0}(2), a), z_3) \} \]

**STEP 2:** One player has one decision node, the other has two decision nodes

Initial states of agent 1 and 2 are \( \pi_{x_0}(1) \) and \( \pi_{x_0}(2) \), respectively. This implies that taking action \( a \) (b) will cause agent 1 to have state of mind \( \Psi(\pi_{x_0}(1), a) \) at \( x_1 \) (\( \Psi(\pi_{x_0}(1), b) \) at \( z_1 \)). The initial state of 1, \( \pi_{x_0}(1) \), also implies that according to agent 1, agent 2 has state \( l_1(\pi_{x_0}(1)) = \pi_{x_0}^1(2) \) and has utility function \( u(\pi_{x_0}^1(2), .) \) at node \( x_0 \).

Agent 1 will make backwards induction as follows: if node \( x_2 \) is reached, he will be in state \( \Psi(\Psi(\pi_{x_0}(1), a), c) \). However, this assessment of agent 1 is relevant only if agent 1 believes that agent 2 will choose action c at node \( x_1 \). To see this, agent 1 has to conjecture about what agent 2 thinks what agent 1 will do at node \( x_2 \). Agent 1 believes that agent 2 will be in state \( \Psi((l(\pi_{x_0}(1)), a) \) at node \( x_1 \). Then agent 1 believes that agent 2 believes that agent 1 will be at state \( \Psi(l(\Psi((l(\pi_{x_0}(1)), a)), c) \).
Figure 1.7. One player has one decision node, the other has two decision nodes.

at node $x_2$. So, agent 1 believes that agent 2 believes that agent 1 will choose $z_3$ if

$$u(\Psi(l(\Psi(l(\pi_{x_0}(1)), a)), c), z_3) \geq u(\Psi(l(\Psi(l(\pi_{x_0}(1)), a)), c), z_4)$$

at node $x_2$ and vice versa. Then, agent 1 believes that agent 2 will compare the chosen terminal node and $z_2$ given agent 2’s conjectured state according to agent 1, which is $\Psi(l(\pi_{x_0}(1)), a)$. If agent 1 believes that agent 2 will choose $z_2$, then agent 1 compares $z_1$ and $z_2$ given his initial state $\pi_{x_0}(1)$. If agent 1 believes that agent 2 will not choose $z_2$, then agent 1 compares $z_1$ and what he thinks he will actually choose at $x_2$ ($z_3$ or $z_4$) given his initial state $\pi_{x_0}(1)$. What he thinks he will actually choose at $x_2$ is $z_3$ if

$$u(\Psi(\pi_{x_0}(1), a), c), z_3) \geq u(\Psi(\pi_{x_0}(1), a), c), z_4)$$

and vice versa.

We can formalize what is mentioned in the last paragraph as follows and this actually gives the definition of the equilibrium:

**Definition 5:** A noncooperative equilibrium of the game with one player having
one and the other having two decision nodes, defined by the set of states, \(\Pi\), the set of actions, \(A\), state dependent utility function, \(u(.,.)\), the transition function, \(\Psi(.,.)\), the function \(l(.)\) and the true transition function, \(\overline{\Psi}(.,.)\), is a path from the initial node to a terminal node such that the initial states are \(\pi_{x_0}(i)\) at the initial node \(x_0\) and the terminal node is given by

\[
(z^{x_0})^* = \arg\max\{ u(\pi_{x_0}(1), z_1), \ u(\pi_{x_0}(1), \widehat{z}) \}
\]

\[
z^{x_1} = \arg\max\{ u(\Psi(l(\pi_{x_0}(1)), a), z_2), u(\Psi(l(\pi_{x_0}(1)), a), \arg\max\{ u(\Psi(l(\Psi((l(\pi_{x_0}(1)), a)), c), z_3), u(\Psi(l(\Psi((l(\pi_{x_0}(1)), a)), c), z_4)\})\})\}
\]

\[
z^{x_2} = \arg\max\{ u(\Psi(\Psi(\pi_{x_0}(1), a), c), z_3), u(\Psi(\Psi(\pi_{x_0}(1), a), c), z_4)\}
\]

\[
If \ z^{x_1} = z_2, \ then \ \widehat{z} = z^{x_1}
\]

\[
If \ z^{x_1} \neq z_2, \ then \ \widehat{z} = z^{x_2}
\]

The actual path in the game does not have to be the same with the initially perceived one. The actual path will be the collection of actions such that

* Actual and conjectured states at each node are determined based on the above specifications.

* At each decision node \(x \in X\), agents choose the action \(a^* \in a(x)\) associated with \((z^x)^*\) where \((z^{x_0})^*\) is defined above; if \((z^{x_0})^* \neq z_1\) or \(x_1\) is reached, then the action \(a^* \in a(x_1)\) associated with \((z^{x_1})^*\) where \((z^{x_1})^*\) is defined as

\[
(z^{x_1})^* = \arg\max\{ u(\overline{\Psi}(\pi_{x_0}(2), a), z_2), u(\overline{\Psi}(\pi_{x_0}(2), a), \arg\max\{ u(\overline{\Psi}(l(\overline{\Psi}(\pi_{x_0}(2), a)), c), z_3), u(\overline{\Psi}(l(\overline{\Psi}(\pi_{x_0}(2), a)), c), z_4)\})\}
\]

If \((z^{x_1})^* \neq z_2\) or \(x_2\) is reached, then the action \(a^* \in a(x_2)\) associated with \((z^{x_2})^*\) where \((z^{x_2})^*\) is defined as

\[
(z^{x_2})^* = \arg\max\{ u(\overline{\Psi}(\pi_{x_0}(1), a), z_3), u(\overline{\Psi}(\pi_{x_0}(1), a), z_4)\}
\]

**STEP 3:** Each player has two decision nodes
This step is a little more complicated because there are different cases that may arise. To define the equilibrium, we have to analyze what each agent believes about how the game will evolve by keeping in mind that the game may actually evolve differently. In the previous step (step 2), player 1 looked ahead and put himself in the place of player 2 and conjectured about what player 2 will choose. If player 1 conjectures that he will choose \( d \) and end the game, then player 1 compares terminal nodes \( z_1 \) and \( z_2 \). If player 1 conjectures that player 2 will choose \( c \), then player 1 compares \( z_1 \) and the terminal node that he thinks he will actually choose at node \( x_2 \) (\( z_3 \) or \( z_4 \)). This is a similar argument to forward induction.

In this step, the arguments are more complicated because there are more cases to analyze. For each case in step 2, there are two corresponding cases in this step that makes total four cases. There are two cases for both of the followings, if player 1 thinks that player 2 thinks that player 1 will choose \( f \) and not choose \( f \) at node \( x_2 \). For the first case, player 1 thinks that the optimal thing to do for player 2 is to choose \( c \) at node \( x_1 \) by expecting player 1 to choose \( f \) after that or to choose \( d \) at node \( x_1 \) (these two refer to case 1 and 2 below, respectively.) For the second case, player 1 thinks that the optimal thing to do for player 2 is to choose \( d \) at node \( x_1 \)
or choose \( c \) at node \( x_1 \) by expecting player 1 not to play \( f \) at node \( x_2 \) (these two refer to case 3 and 4 below, respectively.)

In the light of the above argument, we can write the formal definition of the equilibrium for the above game as follows:

**Definition 6:** A noncooperative equilibrium of the game with both players having two decision nodes, defined by the set of states, \( \Pi \), the set of actions, \( A \), state dependent utility function, \( u(\cdot, \cdot) \), the transition function, \( \Psi(\cdot, \cdot) \), the function \( l(\cdot) \) and the true transition function, \( \overline{\Psi}(\cdot, \cdot) \), is a path from the initial node to a terminal node such that the initial states are \( \pi_{x_0}(i) \) at the initial node \( x_0 \) and the terminal node is given by

\[
(z^{x_0})^* = \text{arg max} \{ u(\pi_{x_0}(1), z_1), u(\pi_{x_0}(1), z^*) \}
\]

\[
z^{x_1} = \text{arg max} \{ u(\Psi(l(\pi_{x_0}(1)), a), z_2), u(\Psi(l(\pi_{x_0}(1)), a), z^{x_2}) \}
\]

\[
z^{x_2} = \text{arg max} \{ u(\Psi(l(\Psi(l(\pi_{x_0}(1)), a)), c), z_3), u(\Psi(l(\Psi(l(\pi_{x_0}(1)), a)), c), z_3), u(\Psi(l(\Psi(l(l(\pi_{x_0}(1)), a)), c)), e), z_4), u(\Psi(l(\Psi(l(l(\pi_{x_0}(1)), a)), c)), e), z_5) \}
\]

\[
z^{x_1'} = \text{arg max} \{ u(\Psi(l(\pi_{x_0}(1)), a), z_2), u(\Psi(l(\pi_{x_0}(1)), a), z_3), u(\Psi(l(\pi_{x_0}(1)), a), c), \arg max \{ u(\Psi(l(\Psi(l(\pi_{x_0}(1)), a)), c), e), z_4), u(\Psi(l(\Psi(l(l(\pi_{x_0}(1)), a)), c)), e), z_5) \} \}
\]

\[
z^{x_2'} = \text{arg max} \{ u(\Psi(l(\pi_{x_0}(1)), a), c), z_3), u(\Psi(l(\pi_{x_0}(1)), a), c), \arg max \{ u(\Psi(l(\Psi(l(\pi_{x_0}(1)), a)), e), z_4), u(\Psi(l(\Psi(l(l(\pi_{x_0}(1)), a)), c)), e), z_5) \} \}
\]

1. If \( z^{x_2} = z_3 \) and \( z^{x_1} = z_3 \), then \( \hat{z} = z^{x_2'} \)
2. If \( z^{x_2} = z_3 \) and \( z^{x_1} = z_2 \), then \( \hat{z} = z^{x_1} \)
3. If \( z^{x_2} \neq z_3 \) and \( z^{x_1'} = z_2 \), then \( \hat{z} = z^{x_1'} \)
4. If $z^{x_2} \neq z_3$ and $z^{x'_1} \neq z_2$, then $\tilde{z} = z^{x'_2}$

Again the actual path may not be the same with the above one. If $(z^{x_0})^* = z_1$, then the game ends immediately. However, if $(z^{x_0})^* \neq z_1$, then the actual path can be found easily by changing the initial state in step 2 accordingly.

The generalization of the equilibrium concept for two player games as we did in one player case needs a very careful examination at different levels of the game tree and remains to be done. Not having a general form of the equilibrium concept for two player games does not prevent us from generalizing it in the frameworks where the states imply simple state transitions and trivial assignments of states to opponents. This is actually the case in so called "Naive Backwards Induction" solution concept that will be explained in the next section and it is a special case of the equilibrium concept presented here. Since the paper by Akin (2005) uses this solution concept, the analysis in the next section should be enough for our purposes.

1.3 A Special Case of Time Inconsistent Preferences

In this part, we will illustrate how the quasi-hyperbolic model (Laibson, 1997; Phelps-Pollak, 1968; Strotz, 1956) fits in the constructed framework. The preference structure is as follows: there are possibly four types of players: exponential (EA), Naive hyperbolic (NHA), Sophisticated hyperbolic (SHA) and partially naive hyperbolic (PNHA). The EA has the following sequence of discount factors: 

$\{1, \delta, \delta^2, \delta^3, \ldots\}$. The NHA, the SHA and the PNHA all have the following same sequence of discount factors: 

$\{1, \beta \delta, \beta \delta^2, \beta \delta^3, \ldots\}$ where $\delta$ is the standard time consistent impatience with $\delta \in (0, 1)$, $\beta$ is time inconsistent preference for immediate gratification or the self-control problem of the agent with $\beta \in (0, 1)$. $\hat{\beta}$ represents a person’s belief about his future self-control problems. In other words, $\hat{\beta}$ is his belief about what his $\beta$ will be in all future periods. The NHA believes he will not have self-control problems in the future, therefore, $\hat{\beta} = 1$. The SHA knows exactly what future self-control problems will be, therefore, $\hat{\beta} = \beta$. The partially naive person has perceptions $\hat{\beta} \in (\beta, 1)$.
In the state of mind specification, there are more types than the ones specified above. The above specification does not take second order beliefs as part of the type. However, in our specification, there may be different naive types, different exponential types and so on. For our purposes, we introduce first the exponential, naive and sophisticated states as follows:

\[
\pi(N) \quad \Rightarrow \quad u(\pi(N), w^t) = w_0 + \beta \sum_{i=1}^{t} \delta^i w_i \text{ where } w^t = \{w_j\}_{j=0}^{t} \\
\Rightarrow \quad \Psi(\pi(N), a) = \Psi(\pi(N), a) = \pi(N) \forall a \in A \text{ that occurs in the same period} \\
\Rightarrow \quad \Psi(\pi(N), a) = \pi(E_2); \quad \Psi(\pi(N), a) = \pi(N) \forall a \in A \text{ that leads to the next period} \\
\Rightarrow \quad l(\pi(N)) = \pi(E_1)
\]

\[
\pi(E_1) \quad \Rightarrow \quad u(\pi(E_1), w^t) = w_0 + \sum_{i=1}^{t} \delta^i w_i \text{ where } w^t = \{w_j\}_{j=0}^{t} \\
\Rightarrow \quad \Psi(\pi(E_1), a) = \Psi(\pi(E_1), a) = \pi(E_2) \forall a \in A \\
\Rightarrow \quad l(\pi(E_1)) = \pi(N)
\]

\[
\pi(E_2) \quad \Rightarrow \quad u(\pi(E_2), w^t) = w_0 + \sum_{i=1}^{t} \delta^i w_i \text{ where } w^t = \{w_j\}_{j=0}^{t} \\
\Rightarrow \quad \Psi(\pi(E_2), a) = \Psi(\pi(E_2), a) = \pi(E_2) \forall a \in A \\
\Rightarrow \quad l(\pi(E_2)) = \pi(E_2)
\]

\[
\pi(\overline{E}) \quad \Rightarrow \quad u(\pi(\overline{E}), w^t) = w_0 + \sum_{i=1}^{t} \delta^i w_i \text{ where } w^t = \{w_j\}_{j=0}^{t} \\
\Rightarrow \quad \Psi(\pi(\overline{E}), a) = \Psi(\pi(\overline{E}), a) = \pi(\overline{E}) \forall a \in A \\
\Rightarrow \quad l(\pi(\overline{E})) = \pi(N)
\]
\[ \pi(S) \Rightarrow u(\pi(S), w^t) = w_0 + \beta \sum_{1}^t \delta^i w_i \text{ where } w^t = \{w_j\}_{j=0}^{t} \]
\[ \Rightarrow \Psi(\pi(S), a) = \overline{\Psi}(\pi(S), a) = \pi(S) \quad \forall \ a \in A \]
\[ \Rightarrow l(\pi(S)) = \pi(\hat{E}) \]

\[ \pi(\hat{E}) \Rightarrow u(\pi(\hat{E}), w^t) = w_0 + \Sigma_{i=1}^t \delta^i w_i \text{ where } w^t = \{w_j\}_{j=0}^{t} \]
\[ \Rightarrow \Psi(\pi(\hat{E}), a) = \overline{\Psi}(\pi(\hat{E}), a) = \pi(\hat{E}) \quad \forall \ a \in A \]
\[ \Rightarrow l(\pi(\hat{E})) = \pi(S) \]

The agent with initial state \( \pi(N) \) is naive as defined at the beginning of the section. He believes that for any action taken during the game, he will move to a kind of an exponential state, \( \pi(E_2) \) and he believes that his opponent is another kind of an exponential agent with state \( \pi(E_1) \). The exponential agent whose type is \( \pi(E_1) \) is exponential in his discounting. He believes that he will move to another kind of an exponential state, \( \pi(E_2) \) for any action and he believes that his opponent’s state is \( \pi(N) \). The exponential agent whose type is \( \pi(E_2) \) is exponential in his discounting. He believes that he will stay in the same state, \( \pi(E_2) \) for any action and he believes that his opponent’s state is same with himself. The exponential agent whose type is \( \pi(E) \) is the “actual” exponential agent who believes that he will stay in the same state, \( \pi(E) \) for any action and believes that his opponent’s state is \( \pi(N) \).

On the other hand, the agent with initial state \( \pi(S) \) is sophisticated as defined at the beginning of the section. He believes that for any action taken during the game, he will stay in the same state \( \pi(S) \) and he also believes that his opponent is an exponential agent with state \( \pi(\hat{E}) \). The exponential agent whose type is \( \pi(\hat{E}) \) believes that he will stay in the same state, \( \pi(\hat{E}) \) for any action and he believes that his opponent is sophisticated, \( \pi(S) \). Thus, the states are common knowledge in the game between these two types.

Now, we introduce the partially naive state as follows:
\[ \pi(P_{N,m,n}) \]
\[ \Rightarrow u(\pi(P_{N,m,n}), w^t) = w_0 + \beta \sum_{i=1}^{t} \delta^i w_i \text{ where } w^t = \{w_j\}_{j=0}^{j=t} \]
\[ \Rightarrow \Psi(\pi(P_{N,m,n}), a) = \overline{\Psi}(\pi(P_{N,m,n}), a) = \pi(P_{N,m,n}) \]
\[ \forall a \in A \text{ that occurs in the same period} \]
\[ \Rightarrow \Psi(\pi(P_{N,m,n}), a) = \pi(P_{E,m,n}); \overline{\Psi}(\pi(P_{N,m,n}), a) = \pi(P_{N,m,n}) \]
\[ \forall a \in A \text{ that lead to the next period} \]
\[ \Rightarrow l(\pi(P_{N,m,n})) = \pi(P_{E_1}) \]

\[ \pi(P_{E,m,n}) \]
\[ \Rightarrow u(\pi(P_{E,m,n}), w^t) = w_0 + \tilde{\beta}_{m,n} \sum_{i=1}^{t} \delta^i w_i \]
where \( w^t = \{w_j\}_{j=0}^{j=t}, \tilde{\beta}_{m,n} = (1 - (1 - \beta) \frac{m}{n}), 0 \leq m \leq n = 1, 2... \)
\[ \Rightarrow \Psi(\pi(P_{E,m,n}), a) = \overline{\Psi}(\pi(P_{E,m,n}), a) = \pi(P_{E,m,n}) \forall a \in A \]
\[ \Rightarrow l(\pi(P_{E,m,n})) = \pi(P_{E_2}) \]

\[ \pi(P_{E_1}) \Rightarrow u(\pi(P_{E_1}), w^t) = w_0 + \sum_{i=1}^{t} \delta^i w_i \text{ where } w^t = \{w_j\}_{j=0}^{j=t} \]
\[ \Rightarrow \Psi(\pi(P_{E_1}), a) = \overline{\Psi}(\pi(P_{E_1}), a) = \pi(P_{E_2}) \forall a \in A \]
\[ \Rightarrow l(\pi(P_{E_1})) = \pi(P_{N,m,n}) \]

\[ \pi(P_{E_2}) \Rightarrow u(\pi(P_{E_2}), w^t) = w_0 + \sum_{i=1}^{t} \delta^i w_i \text{ where } w^t = \{w_j\}_{j=0}^{j=t} \]
\[ \Rightarrow \Psi(\pi(P_{E_2}), a) = \overline{\Psi}(\pi(P_{E_2}), a) = \pi(P_{E_2}) \forall a \in A \]
\[ \Rightarrow l(\pi(P_{E_2})) = \pi(P_{E,m,n}) \]

\[ \pi(\overline{P_{E}}) \Rightarrow u(\pi(\overline{P_{E}}), w^t) = w_0 + \sum_{i=1}^{t} \delta^i w_i \text{ where } w^t = \{w_j\}_{j=0}^{j=t} \]
\[ \Rightarrow \Psi(\pi(\overline{P_{E}}), a) = \overline{\Psi}(\pi(\overline{P_{E}}), a) = \pi(\overline{P_{E}}) \forall a \in A \]
⇒ \( l(\pi(PE)) = \pi(PN_{m,n}) \)

The agent with state \( \pi(PN_{m,n}) \) is partially naive in the sense that he believes that he will have a self-control problem \( \hat{\beta} \) in all the future periods. This is represented as he thinks he will move to another state \( \pi(PE_{m,n}) \) that implies a self control problem \( \hat{\beta} \) and he will stay at this state for any taken action\(^3\). He also thinks that his opponent is an exponential agent, \( l(\pi(PN_{m,n})) = \pi(PE_1) \) and he thinks his opponent shares the same beliefs with himself, \( l(\pi(PE_1)) = \pi(PN_{m,n}) \).

Moreover, after he moves to the state \( \pi(PE_{m,n}) \), his opponent will move to another state too, \( \pi(PE_2) \), at which he shares again the same beliefs with his opponent, \( l(\pi(PE_{m,n})) = \pi(PE_2) \) and \( l(\pi(PE_2)) = \pi(PE_{m,n}) \). The agent with state \( \pi(PE) \) is an exponential agent who knows his opponent is in the partially naive state, \( l(\pi(PE)) = \pi(PN_{m,n}) \).

The equilibrium concept that we defined in section 2 would represent the "Naive Backwards Induction" (NBI) if the game is played between the naive type with state \( \pi(N) \) (or partially naive with state \( \pi(PN_{m,n}) \)) and the exponential type with state \( \pi(\overline{E}) \) (or \( \pi(\overline{PE}) \)). Now we will give an example of how this is represented in our solution concept. Take the example game with one player having one and the other having two decision nodes with a small difference.

The only difference is that the outcome \( w_3 \) is earned not on date 2 but on date 1. The first player has initial state \( \pi_{x_0}(1) = \pi(N) \) and the second player has initial state \( \pi(\overline{E}) \). We suppose that the outcomes and the discount factors are such that:

\[
\begin{align*}
w_{11} &< \beta \delta w_{21} \\
\delta w_{42} &< w_{22} < w_{32} \\
\beta \delta w_{41} &< w_{31} < \delta w_{41}
\end{align*}
\]

We now apply the equilibrium concept. Player 1 will choose the associated action

\(^3\)Specifying \( m \) and \( n \) in a certain way leads the partially naive state to become either completely naive or sophisticated states defined earlier. The case where \( m = 0 \) refers to the naive case where \( \pi(PE_{m,n}) = \pi(PE_2) \). The case where \( m = n \) refers to the sophisticated case where \( \pi(PN_{m,n}) = \pi(PE_{m,n}) \) and \( \pi(PE_1) = \pi(PE_2) = \pi(\overline{PE}) \).
Figure 1.9. An example for naive backwards induction.

with \((z^{x0})^*\) given by this.

\[
(z^{x0})^* = \arg \max \{ u(\pi_{x0}(1), z_1), u(\pi_{x0}(1), \hat{z}) \}
\]

\[
z^{x1} = \arg \max \{ u(\Psi(l(\pi_{x0}(1)), a), z_2), u(\Psi(l(\pi_{x0}(1)), a), \arg \max \{ u(\Psi(l(\psi((l(\pi_{x0}(1)), a)), c), z_3), u(\Psi(l(\psi((l(\pi_{x0}(1)), a)), c), z_4)) \}) \}
\]

\[
z^{x2} = \arg \max \{ u(\Psi(\pi_{x0}(1), a), c), z_3), u(\Psi(\pi_{x0}(1), a), c), z_4) \}
\]

\[
If \ z^{x1} = z_2, \ then \ \hat{z} = z^{x1}
\]

\[
If \ z^{x1} \neq z_2, \ then \ \hat{z} = z^{x2}
\]

We first find \(z^{x1}\). The following argument results in choosing \(f\) at node \(x_2\).

\[
\arg \max \{ u(\pi(E_2), z_3), u(\pi(E_2), z_4) \} = \arg \max \{ w_{31}, \delta w_{41} \} = z_4 \Rightarrow \text{action } f
\]

Then \(z_2\) and \(z_4\) will be compared by agent 2 and since the following is satisfied,
\[ z^{x_1} = z_2 \text{ and action } d \text{ will be chosen at node } x_1 : \]
\[ z^{x_1} = \arg \max \{u(\pi(E_2), z_2), \ u(\pi(E_2), z_4)\} \]
\[ z^{x_1} = \arg \max \{w_{22}, \ \delta w_{42}\} = z_2 \Rightarrow \text{ action } d \]

Then, since \[ z^{x_1} = z_2 \text{ and } (z^{x_0})^* = \arg \max \{u(\pi(N), z_1), \ u(\pi(N), z_2)\} = z_2, \text{ Player 1 will choose action } a \text{ at node } x_0. \]

What actually happens after action \(a\) is chosen is that player 2 will choose \((z^{x_1})^*\) where

\[ (z^{x_1})^* = \arg \max \{u(\Psi(\pi(E), a), z_2), \ u(\Psi(\pi(E), a), \ \arg \max \{u(\Psi(l(\Psi(\pi(E), a)), c), z_3), u(\Psi(l(\Psi(\pi(E), a)), c), z_4)\})\} \]

\[ (z^{x_1})^* = \arg \max \{u(\pi(E), z_2), \ u(\pi(E), \ \arg \max \{u(\pi(N), z_3), u(\pi(N), z_4)\})\} \]

Since \(\Psi(l(\Psi(\pi(E), a)), c) = \Psi(l(\pi(E)), c) = \Psi(\pi(N), c) = \pi(N). \) This is true because action \(c\) is taken in the same period and does not lead to any change in the state of the first player.

\[ (z^{x_1})^* = \arg \max \{u(\pi(E), z_2), \ u(\pi(E), \ \arg \max \{w_{31}, \ \beta \delta w_{41}\})\} \]

\[ = \arg \max \{u(\pi(E), z_2), \ u(\pi(E), z_3)\} = z_3 \]

Action \(c\) is chosen since \((z^{x_1})^* = z_3. \) Afterwards, the action \(a^* \in a(x_2)\) will be chosen associated with \((z^{x_2})^*\) where \((z^{x_2})^*\) is defined as

\[ (z^{x_2})^* = \arg \max \{u(\Psi(\pi(N), a), c), \ u(\Psi(\pi(N), a), c), z_4)\} \]

\[ = \arg \max \{u(\pi(N), z_3), \ u(\pi(N), z_4)\} \]

Action \(e\) is chosen because \((z^{x_2})^* = z_3. \) From the perspective of date 0, for agent 1, strategy \((a, f; d)\) seemed to be the optimal strategy but the actual play
turned out to be \((a, e; c)\).

1.4 Discussion

In this paper, the equilibrium concept for perfect information extensive form games is defined as a path from the initial node to a terminal node given the initial state of the first mover. When this is done, all the evaluations are based on the initial state of player 1, the first mover. This is because, without loss of generality, player 1 starts the game and his exogenously given initial state determines his own and his opponent’s valuations of each terminal node since the opponent’s conjectured state is also determined by his initial state. Since the states embrace, almost, all the information structure of the game, by definition, and player 1 is the first mover, everything is driven by player 1’s initial state, \(\pi_{x_0}(1)\). However, this is just for the first move. After the first move made by the first player, the true transitions lead to a potentially different states for each agent, all the evaluations potentially change and now all evaluations are based on what the second player’s true state of mind at the node reached is. However, this is true for the second move only and so on. Thus, there is a dynamic structure in this framework where the set of strategies is the same but the optimal strategies of the players potentially change during the course of the game.

A strategy in general tells a player what to do at each of his information set. The optimal strategy or the equilibrium strategy is, in general, obtained by a best response argument. In this framework, it works in a similar way. Each player has a set of strategies (in a perfect information game like this, it is basically the Cartesian product of the set of actions at each of his nodes.) Then, each player (given his state of mind that implies a state for his opponent, assigns a real number to each terminal node and implies a specific sequence of states corresponding to each node that depends on the actions chosen) determines the optimal action for himself and for his opponent at each node based on his given state. At each node, the optimal strategy may vary for both agents depending on how their states evolve. In this sense, the equilibrium concept is dynamic but also deterministic.

In the existing framework, given the states, the equilibrium boils down to a deterministic payoff maximization. Each player has point conjectures about his own
future states (characterized by $\Psi(.,.)$) and their opponent’s state (characterized by $l(.)$) but both of these may turn out to be wrong. Instead of this, a probabilistic approach can be introduced in such a way that both of these may vary stochastically (both functions $\Psi(.,.)$ and $l(.)$ assign states probabilistically.)

1.5 Conclusion

In this paper, I study an equilibrium concept for both one and two player extensive form games. The players are modeled as agents being at potentially different states of minds. The state space is inclusive and comprehensive enough such that it embraces almost all the information structure of the game and the all possible types of the players. By this, the aim is to introduce a general enough framework in order to cover different bounded rationality approaches. Finally, a special case that studies the games played by potentially time inconsistent agents is presented. This paper studies the equilibrium concept in perfect information, one and two player extensive form games. As extensions, it can be generalized to other classes of extensive form games and different psychological, behavioral and other bounded rationality models can be illustrated in a formal way such as regret, limited memory, limited foresight, present bias in a more general sense and addiction.

1.6 References


Chapter 2

TIME INCONSISTENCY AND LEARNING IN BARGAINING GAMES

2.1 Introduction

In our daily lives, we often face decisions for alternative actions to be made over time. Traditional economic analysis assumes that people are rational, and they take actions that maximize their payoff or utility. It also assumes that rational people behave consistently, which means they follow the original contingent plan or strategy, when they make these decisions. In other words, a rational agent’s goals at different dates cannot be in conflict, and he always agrees with his future selves. This means decisions will not have to follow a sort of self-conversation: "OK, I had decided to do this before; but now, let me do something else." (e.g., "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 their vulnerability 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. Quasi-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). From here on, hyperbolic discounting means only quasi-hyperbolic discounting. Specifically, it captures the sharp short-
run drop in valuation observed in the experimental studies, and it is analytically tractable. It also captures the key observation that there is a faster rate of decline in the short-run than in the long-run.

Standard economic models assume that people are exponential discounters, and they are able to design time consistent plans. In this sense, exponential discounters are rational. On the other hand, the literature on time inconsistency (also known as preference reversals and self-control problems) introduced naive, partially naive and sophisticated types of agents. 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. These agents are keen on their near-future selves’ gratification such that their distant-future selves probably will be regretful about this pursuit of immediate gratification. In this sense, hyperbolic discounters are boundedly rational.

A naive time inconsistent agent procrastinates. However, deadlines and, potentially, learning about their preference reversals prevents this her\(^1\) from procrastinating costly tasks forever. This paper examines the effect of learning on the naive agent’s decisions in a bargaining framework.

In the model considered here, two players -each of whom is one of the following types; naive, sophisticated, partially naive or exponential, are engaged in an infinite horizon alternating-offers bargaining game under the assumption that each player knows the opponent’s type. At each case, the equilibrium of games played between different types of agents, such as naive-sophisticated and exponential-naive, is characterized. The focus will be more on the games involving naive and partially naive agents. During the game, partially naive agents, by observing possible rejections, might gradually become more aware of their naivete. This is modeled similar to Yildiz (2004), although Yildiz examines optimism about recognition process rather than self-control.

The learning framework I assume works as follows: the partially naive agent has an initial belief about her future self-control problems. She interprets the rejections during the course of the game in such a way that her actual self-control

\(^1\)There is no specific reason that some agents are "she" and others are "he". It is just for convenience.
problems might actually be more severe than perceived and she updates her beliefs accordingly. On the other side, the opponent has a trade-off between extracting more rent from the partially naive agent (by making rejected offers) and bearing the cost of delaying. 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. There are two solution concepts, ”naive backward induction” and ”equilibrium,” developed by Sarafidis (2004) for the games played by time inconsistent agents. In the naive backward induction (NBI), the naive agent plays a best response to which the opponent may play. However, due to self-misconceptions and misconceptions about the opponent, surprises may occur during the course of the game. Thus, due to this incorrect belief formation, the naive agent is not able to anticipate the opponent’s actions 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.

In this paper, since the assumptions of the NBI better fits to the framework here and the foundations of the NBI and behavioral assumptions about the players match better, I will focus more on the NBI solution concept. The following interesting results arise from using NBI solution concept. Offers of the players who are naive or partially naive are never accepted by any type of player in both no learning and gradual learning cases. I show that without learning, there is immediate agreement between the naive or partially naive agent and the exponential agent when the exponential agent offers first. When the naive agent offers, there is a one period delay. On the other hand, two naive agents perpetually disagree. Moreover, a player’s share increases with the degree of naivete. Second, I consider naive agents who can learn their types over time and show that there is a critical date before which no agreement occurs.

For each case, where different types of agents are engaged in a sequential bargaining game, I examine whether a delay exists and what the equilibrium shares are. Of the results mentioned above, the two most important ones are: First, the existence of time inconsistent agents who may learn to be more consistent over time may explain bargaining delays to some extent and second, the unawareness
of the agent about her time inconsistency makes her better off except in a case when playing against the same type of agent.

There is a growing literature about time inconsistency\textsuperscript{2}. First, Strotz (1956) proposed that individuals may not have stationary preferences over time. They might value close satisfactions more than distant ones. Phelps and Pollak (1968) formalized this kind of behavior in a mathematically more convenient way, called the $\beta \delta$ approach. Afterwards, Laibson (1997) used this formalization to explain individuals’ observed saving behaviors. O’Donoghue and Rabin (1999a, 1999b and 2001) examined time inconsistent individuals’ decision making in different economic environments such as principal-agent problems, when to complete a task, which option to choose from a menu of options and when to choose. Here, this formalization is considered in the context of a non-cooperative game.

We use the alternating-offers bargaining framework proposed first by Rubinstein (1982). Rubinstein assumed that preferences are stationary 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. The model in this study has a non-stationary environment as well but this comes directly from the preferences of the players.

Akin (2004) considered an alternating-offers bargaining game between two players. One of the players is time inconsistent and cannot learn about self-preferences. Bargaining, in that paper, is the second stage of a two stage principal-agent game whose first stage is self-investment of the agent. On the other hand, this paper introduces learning in a sequential bargaining context.

Although some of the literature just briefly mentions potential learning considerations as extensions of the existing models (see, O’Donoghue and Rabin (2001) and Sarafidis (2004)), there is no distinct work, as far as I know, focusing on this important aspect of time inconsistency. Dellavigna and Malmendier (2003) examined self-control in the market and showed that some of the observed behavior of consumers (for example, in health clubs) can be explained by the time inconsistency of the agents. They also mentioned the effect of these agents’ learning to explain some empirical results:

\textsuperscript{2}See Ainslie (1992) and Loewenstein and Prelec (1992) for an extensive discussion.
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-fifths 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.

To address this issue, I consider not only time inconsistency in the context of non-cooperative games but I also look at the case where time inconsistent agents may learn to be more consistent over time. In order to see the pure effect of time inconsistency and learning, the environment is kept as simple as possible. The assumption is that each agent knows the opponent’s type, and the Bayesian type incomplete information between the agents that allows reputation formation is ruled out. Moreover, any mechanism is ruled out that also leads to delay mentioned in the literature such as more than two players with optimism, stochastic pie size or optimism about the outside option.

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, the focus is on individuals being more sophisticated about their self-control problems during the course of a bargaining game.

The remainder 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. Finally, section 6 concludes the paper with a brief discussion of the results. Some of the proofs are presented in the Appendix.

### 2.2 Solution Concept: Naive Backwards Induction

The solution concept used in this paper depends crucially on the beliefs that agents have about offers and about their future selves. Therefore, the discussion of these beliefs is first. Players may be one of four types: naive, sophisticated, partially naive or exponential. The naive hyperbolic agent (NHA) is not aware of her time
inconsistency, which means she thinks that she will be patient, but, in reality, will be impatient in future periods. The sophisticated hyperbolic agent (SHA) is fully aware of his time inconsistency and behaves accordingly. The partially naive agent is aware of her future self-control problems only to some extent. The exponential agent (EA) is time consistent and fully rational. The naive hyperbolic agent has wrong beliefs about herself, and she believes that the rational agent thinks about her what she thinks about herself. Each agent knows the opponent's type.

For games that include only time consistent players, Subgame Perfect Nash Equilibrium (SPNE) is used as the solution concept. In addition, for games that include only time consistent and sophisticated hyperbolic players, since SHA is fully aware of self-preference reversals, may be treated as a time consistent player with standard impatience $\beta \delta$. This implies that preferences are stationary and beliefs are mutually consistent. Thus, SPNE can be applied. On the other hand, if at least one of the players is NHA, then redefinition and use of a slightly different solution concept is necessary. Sarafidis (2004) proposes "Naive Backward Induction" in games with time inconsistent players.

In an NBI, the player who has self-control problems plays a best response to what is thought to be the opponent's play. One caveat of this concept is that since a time inconsistent player may not implement what has been planned for the future and holds wrong beliefs about both the self and the opponent, this player may be surprised by how players (including the self) play when the game proceeds. This agent also believes that the opponent has the same opinion about what the self-perception is. In an NBI, players form beliefs about how others will play the game by putting themselves in the shoes of the other players. However, this belief formation process leads the NHA to anticipate the opponent's actions incorrectly. The NHA recognizes the other player's time inconsistency, if any, but not the one of the self.

---

3 Types will be defined and explained formally in the next section.

4 A person's naivete in beliefs about self-behavior in the future may be caused by over-optimism or over-confidence or both. Contemporary psychologists agree that "on nearly any dimension that is both subjective and socially desirable, most people see themselves as better than average" (Myers, 1996). Since patience is a kind of ability to endure waiting or delay without becoming annoyed or upset (dictionary definition), it is reasonable to think that people may consider themselves more patient than others. It is argued that the players tend to neglect their own self-serving biases, although they recognize the biases of other players (e.g., Babcock et al. (1995)). The agent may be self-naive because of over-optimism and may be naive about the relative
This solution concept is called naive backwards induction because the naive players work backwards but do this based on their naive beliefs. Backwards induction can only be applied to the finite-horizon games. However, with the quasi-hyperbolic discounters, infinite-horizon games can be reduced to multi-stage finite-horizon games such that in the subgames that the later stages start, the solutions can be found based on players’ beliefs. Solution for the infinite-horizon bargaining games with naive types is as follows: Using Rubinstein’s results for the continuation game, given the beliefs of each agent, essentially reduces the game to a two-stage game, today and the future. This is because each player believes that each of the future self will have the same self control problem, $\hat{\beta}$, although different players may have different $\hat{\beta}$ values. Then, the NBI applies. At time $t$, suppose that the conjectured continuation value for a player given self-beliefs and the other player is $v(t+1)$. The assumption is that continuation values for each player exist. Then, the current period’s action given these continuation payoffs can be solved. Thus, if all the subgames played by different selves of the players have a solution, then one can find the NBI for the whole (finite or infinite) game.

Note that, with fixed types (no learning,) the continuation values do not change over time. However, with learning, they may change during the course of the game. In addition, potentially differing beliefs that each agent holds (about each agent) may lead to different conjectured continuation payoffs. This is actually the main reason for potential delays in agreement.

At this point, construction of the required notations for formal definitions is necessary. Two players play an alternating-offers bargaining game. One player is a time consistent exponential agent and the other is a time inconsistent naive agent. The scenario is to create an infinite sequence of fictitious players from the NHA. Let $NHA_t(i)$ represent the $t-$period self of the agent naive hyperbolic agent $i$, $NHA(i)$. $NHA_t(i)$ has available actions that the $NHA(i)$ has from period $t$ onwards.

Let $s_{EA}$ be a history dependent strategy for the EA and $s_{NHA_t}(i)$ be a history patience in when comparing the self with the others because of positive self-image. (See Brocas and Carrillo (2004) and Zabojnik (2004) for rationalization of positive self image; Pinto and Sobel (2005) for reasons and extensive discussion of positive self image; Camerer and Lovallo (1999) and Hoelzl and Rustichini (2005) for investigation of positive self-image by using economic incentives.)

As mentioned, this solution concept is mainly for games including at least one naive hyperbolic agent. Games with other types of agents are also examined.
dependent strategy of $NHA_t(i)$ or

$$s_{NHA_t(i)} = \{s_{NHA_t(i)}^k\}_{k=t}^{\infty}$$

where $s_{NHA_t(i)}^k$ represents the plan of $NHA_t(i)$ about what she will play at time $k$, $t < k < \infty$.

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

$$s_{NHA(i)} = \{s_{NHA_t(i)}^t\}_{t=0}^{\infty}$$

The above arguments motivate the following formal definitions of the NBI when there is at least one naive agent.

**Definition 1:** 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 $NHA_t$.

**Definition 2:** $s = (s_{NHA_0(1)}, s_{NHA_1(1)}, ...; s_{NHA_0(2)}, s_{NHA_1(2)}, ...)$ is a strategy profile that constitutes Naive Backwards Induction (NBI) solution if each $s_{NHA_t(i)}$ is a best response to $s_{NHA(j)}$ for $NHA_t(i), i \neq j \in \{1, 2\}$.

Each player has conjectures about both their own type and their opponent’s type. The assumption is that each player’s conjecture about the opponent’s type is correct but about the self may not be correct. Given this, each player plays a best response to what is thought to be the opponent’s play.

An example, as in the following figure, can be given to make the solution concept clearer: There are two players one of whom is NHA (Player 1) and the other (Player 2) is either exponential or naive; they play the following game. First, the player 2 is exponential. There are three periods. In the first period, the NHA chooses ”in” or ”out.” In the second period, the NHA chooses ”left” or ”right” after seeing the action of EA, ”Left” or ”Right.” First number in the parenthesis is the payoff of the NHA and second number is the payoff of the EA. Payoffs with superscript star are obtained at the end of the third period. Other payoffs are obtained at the end of the second period. Also assume that $x < 10$. 

For simplicity and to emphasize the effect of time inconsistency, $\delta = 1$ and $\beta = 0.5$ are imposed. The NHA thinks, at $t=1$, that the self-preference will be "8 at $t = 3$" over "7 at $t = 2$" and "10 at $t = 3$" over "9 at $t = 2." However, at $t = 2$, the choice is "7 at $t = 2$" to "8 at $t = 3" and "9 at $t = 2" to "10 at $t = 3" because of the self-tendency for immediate gratification. In other words,

\[
\begin{align*}
\text{at } t &= 1, \quad \text{NHA} \rightarrow 8_{t=3} \succ 7_{t=2} \text{ and } 10_{t=3} \succ 9_{t=2} \\
\text{at } t &= 2, \quad \text{NHA} \rightarrow 8_{t=3} \prec 7_{t=2} \text{ and } 10_{t=3} \prec 9_{t=2}
\end{align*}
\]

**Figure 2.1.** An example of naive backwards induction with two different types of opponent.

The example is shown in Figure 2.1. According to the NBI, the NHA plays best response to what is thought to be the opponent’s play. In the example, at $t = 1$, the NHA thinks that any choice of the EA will result in a "right" choice at $t = 2$. The NHA thinks that the EA shares this belief and that the EA will choose "Left" because $4 > 3$. So, if "in" is chosen, the EA will choose "Left" and the NHA will get 10 at $t = 3$. Thus, the NHA chooses "in." Thus (Left; in, right, right) constitutes an NBI solution. However, the actual play is (Left; in, left, left).
because player 1 changes preferences over the payoffs at $t = 2$.

Now, the player 2 is naive. $NHA(1)$ thinks as explained above. Similarly, $NHA(2)$ thinks at $t=1$ that the choice will be "6 at $t = 3$" over "5 at $t = 2$" and "4 at $t = 3$" over "3 at $t = 2". However, at $t = 2$, the preference will be "5 at $t = 2$" to "6 at $t = 3$" and "3 at $t = 2" to "4 at $t = 3." In other words,

\[
\begin{align*}
ad t = 1, & \quad NHA(2) \to 6_{t=3} \succ 5_{t=2} \land 4_{t=3} \succ 3_{t=2} \\
ad t = 2, & \quad NHA(2) \to 6_{t=3} \prec 5_{t=2} \land 4_{t=3} \prec 3_{t=2}
\end{align*}
\]

We can summarize the beliefs of each agent as follows:

$NHA(1)$:

Player 1 believes that her play will be "right" independent of the node at which she is located at $t = 2$.

Player 1 believes that player 2 will play "Right" even if he thinks that his play will be "Left" because he will change choice at $t = 2$.

$NHA(2)$:

Player 2 believes that his play will be "Left."

Player 2 believes that player 1 will play "left" independent of the node at which she is located because she will change choice at $t = 2$.

In the example, if $x = 9$, then (Right; out, right, right) constitutes the unique NBI because of the following: $NHA(1)$ thinks that $NHA(2)$ will play "Right," then she will play "right" and get 8 at $t = 3$. Since playing "out" gives 9, she plays "out." So, $NHA_1$ plays "out" and the game ends.

If $x = 7$, then (Right; in, right, right) constitutes the unique NBI because of the previous reasoning except now playing "in," she thinks, gives more payoff.

On the other hand, at $t = 1$, $NHA(2)$ thinks that $NHA(1)$ will play "left" independent of the node at which she is located. Then, $NHA(2)$ thinks 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 preference and play the other strategies that they have. Specifically, $NHA(1)$ plays "left" at both her information sets and $NHA(2)$ plays "Right." 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 each thinks the
other player will play. Each agent realizes the other agent’s time inconsistency but not the self-inconsistency. They are right about their opponents but not right about their own actions.

Now, the model is introduced and the equilibrium outcomes are examined in the cases where, first, updating of beliefs is not allowed and second, updating of beliefs is possible.

2.3 Model

Let \( T = \{0, 1, 2, 3, \ldots\} \) 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. \( U \) is the set of feasible utility pairs, \( U = \{u \in [0, 1]^2 | u_1 + u_2 \leq 1\} \). At each time \( t \), \( i \) offers a utility pair \( u = \{u_1, u_2\} \in U \) (payoffs or shares are treated as utilities.) 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.

The players in this model 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: \( \{1, \delta, \delta^2, \delta^3, \ldots\} \). The NHA, the SHA and the partially naive agent have the following sequence of discount factors: \( \{1, \beta \delta, \beta \delta^2, \beta \delta^3, \ldots\} \) where \( \delta \) is the standard time consistent impatience with \( \delta \in (0, 1) \), \( \beta \) is time inconsistent preference for immediate gratification or the self-control problem of the agent with \( \beta \in (0, 1) \). Let \( \hat{\beta} \) be a person’s belief about her future self-control problems. In other words, \( \hat{\beta} \) is her belief about what her taste for immediate gratification, \( \beta \), will be in all future periods. The NHA believes she will not have self-control problems in the future, therefore she has perceptions \( \hat{\beta} = 1 \). The SHA knows exactly what future self-control problems will be, therefore, this player has perceptions \( \hat{\beta} = \beta \). The partially naive person has perceptions \( \hat{\beta} \in (\beta, 1) \).

Naive players are certain at every information set that it is common knowledge that they will not have time inconsistency in the future. The NHA recognizes the other player’s time inconsistency, if any, but not of the self. In addition, if a naive player’s opponent is sophisticated or exponential, then the opponent will recognize the naivete of the player. On the other hand, sophisticated players
certainly believe that at every information set that it is common knowledge that they will have time inconsistency in the future. The assumption is that each type knows the opponent’s type in the bargaining game.

A dynamic preference structure can be imposed on naive and partially naive agents’ beliefs because they are not fully aware of their future preference reversals. Therefore, learning can be incorporated 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. In the first approach, \( \hat{\beta} \in (\beta, 1] \) does not change over time, which means agents believe that their self-control problem, \( \beta \), will disappear (\( \beta \) will be \( \hat{\beta} = 1 \)) or diminish (\( \beta \) will be \( \hat{\beta} \)) after tomorrow, and they will not change this belief whatever happens in the future. They believe that they will discount \( t + 1 \) by \( \hat{\beta}\delta \) at time \( t > 1 \). In the second approach, the assumption is that players learn immediately whenever they observe a rejection (either rejection of their offer or they reject an offer), that is, \( 1 \geq \hat{\beta} > \beta \) becomes \( \hat{\beta} = \beta \) immediately after one rejection. In the third approach, the behavior of players who learn to be more sophisticated gradually in time is examined, that is, \( \hat{\beta} \) may not equal 1 and may evolve over time and gets closer to \( \beta \). They learn by introspection about themselves during the evolution of the game.

The three approaches are examined in order. First, no learning is assumed. Then, the other two approaches allowing learning are considered.

### 2.4 Equilibrium Without Learning

This section studies bargaining games including at least one naive agent by using NBI as the solution concept. Each case where different types of agents are involved is analyzed separately. The assumption is that the NHA never updates self-beliefs or the beliefs about the opponent. The unique solutions found at different cases directly follows from Rubinstein (1982).

In the general framework of the Rubinstein model, players reach an agreement immediately in equilibrium since there is discounting. I assume that strategies are stationary for every player. At every even period, \( t = 0, 2, 4, \ldots \), player 1 offers \( 1 - x^* \) to player 2 such that player 2 accepts and their realized payoffs would be
(x^*, 1 - x^*), respectively. Similarly, at every odd period, \( t = 1, 3, 5 \ldots \), player 2 always offers \( y^* \) to player 1.

An NHA believes with probability one that it is common knowledge that no self time inconsistency will exist in the future. The other player will recognize this, but, without learning, will also recognize that the opponent remains to be time inconsistent. Note that if \( \beta < 1 \) for an agent, this makes her relatively impatient and increases the opponent’s share. If an agent with \( \beta < 1 \) is also naive, then this agent always underestimates the opponent’s reservation value. Thus, a naive player offers the other party less than what the other party would accept, an offer that will be rejected.

On the other hand, an SHA believes with probability one that it is common knowledge that self time inconsistency will arise at each future period. Consequently, the opponent makes offers according to the conjecture that the SHA’s discounting rate is effectively \( \beta \delta \).

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 and have different discount factors \( \delta_1 \) and \( \delta_2 \). The result can be written as either the limit case of the finite horizon game or the recursive way of solving it as in Shaked-Sutton.

**Remark 1:** In the infinite horizon alternating-offers game with both players have exponential discounting with discount factors \( \delta_1 \) and \( \delta_2 \), the equilibrium payoffs are:

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

and \( x^* \) is the payoff of Player 1.

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.

Before examining each case one by one, the following lemma states a general delay result including both no learning and learning cases.\(^6\)

---

\(^6\)In particular, I thank one of the referees for attracting my attention to this extension of the
Lemma 1: Offers of the players with $\beta \in (0, 1)$ and perceptions $\hat{\beta} \in (\beta, 1]$ are never accepted by any type of player in both no learning and gradual learning cases.

The lemma 1 is very general and strong in the sense that offers of both naive and partially naive players will never be accepted by any other type of the player. Also this implies that even with the imposed learning framework, naive agents cannot agree with naive or partially naive agents and similarly partially naive agents cannot agree with partially naive or naive agents.

Without learning, the rationale behind this is the following assumption: each naive and partially agent recognizes the other player’s time inconsistency completely but not fully that of the self. Since there is no learning, players do not update their beliefs nor change their strategies but rather they stay naive no matter how the game evolves. Each of them is so optimistic about their future shares and so obstinate that they insist on offering the same rejected share to the other.

The gradual learning model does not allow the agents’ fully learning about their preferences. Learning slows down overtime and they never completely learn their exact preferences. There is discontinuity in the result when $\hat{\beta} = \beta$ in the sense that for any $\hat{\beta} \in (\beta, 1]$, this result is obtained but when $\hat{\beta} = \beta$, there is either immediate agreement (with exponential, sophisticated, not learning naive and partially naive opponents) or a finite period of delay (with learning naive and partially naive opponents.) Even with very little naivete ($\hat{\beta}$ is very close to $\beta$), no agreement result is obtained. The rationale is as follows: The naive and partially naive agents always offer the present value of what they think their opponent expects to earn by rejecting. However, since they hold different beliefs than the opponent, it turns out that from the perspective of the opponent, rejecting based on these beliefs is always optimal. When there is learning, the following happens: 1. They will update their beliefs in case of a rejection, 2. Their opponent makes them a little more sophisticated that allow the opponents to extract more from them at the next period and 3. Naive players are not aware of their learning process.\(^7\) Because of these reasons, the SHA’s and EA’s expectations turn out to

---

\(^7\)In this context, the naive players may be too naive. They are not even aware that they will update their beliefs. It seems odd to assume a naive player could learn but never be aware that learning will take place. This is analogous to a person who is capable of changing beliefs but also thinks whatever is believed currently is true. An alternative formulation for more “sophisticated” naive players is considered later.
be always higher than what is predicted and offered by naive players.

The formal proof of the lemma using a specific notation appears in the appendix and examines each of these cases. The following result is a corollary of the lemma.

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

The corollary states that naive backward induction gives a disagreement result when two naive agents play the sequential bargaining game. This 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 such that they end up with zero payoff.

Yildiz (2003) analyzes a sequential bargaining framework where players hold subjective and possibly optimistic beliefs about the recognition process. He assumes that the players’ beliefs are common knowledge and the players know the strategies of the other players. However, in this study’s context, naive players don’t know the other player’s strategies. This is the main distinction. Yildiz also points out that if it is common knowledge that the players will remain sufficiently optimistic for a sufficiently long future, they will agree immediately in equilibrium. That is, excessive optimism alone cannot be a reason for a delay in agreement. In the present case, players stay optimistic forever about their own preferences, but no immediate agreement or a delayed agreement occurs. This is again caused by the lack of common knowledge of beliefs.

Proof. The proof of this corollary follows directly from the Lemma 1 where $\beta_i < \hat{\beta}_{ji} = \hat{\beta}_i = 1$ and $\beta_j < \hat{\beta}_{ij} = \hat{\beta}_j = 1$.

The results of the other cases are mentioned in the following propositions.

Proposition 1: Let an EA and an NHA play the alternating-offers bargaining game, then: 1. If the EA is the first proposer, then there is no delay and 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 the NHA 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. The counter offer is as in the previous case, and the game ends with the implied shares at $t = 1$.

The above result indicates that if the naive player stays naive in the entire game, the exponential player recognizes this, makes an offer that will be accepted
and the game ends at the first date that the exponential agent makes an offer.

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 possible for the NHA, the offer is accepted. The NHA will reject any offer less than this because she anticipates to get $1 - y_1 = \frac{1}{1 + \delta}$ tomorrow. 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 the static belief and behavior of the NHA. 2. This follows from the Lemma 1 where $\beta_i < \hat{\beta}_{ji} = \hat{\beta}_i = 1$ and $\beta_j = \hat{\beta}_{ij} = \hat{\beta}_j = 1$.

Proposition 2: Let an SHA and an NHA play the alternating-offers bargaining game, then: 1. If the SHA is the first proposer, then there is no delay and equilibrium shares are $(x^*, 1 - x^*) = (1 - \beta \delta \frac{1 - \beta \delta}{1 - \beta \delta}, \beta \delta \frac{1 - \beta \delta}{1 - \beta \delta})$ where $x^*$ is the share of the SHA. 2. If the NHA is the first proposer, then the NHA offers $(x^*, 1 - x^*) = (1 - \beta \delta \frac{1 - \delta}{1 - \beta \delta}, \beta \delta \frac{1 - \delta}{1 - \beta \delta})$ where $x^*$ is the share of the NHA. The SHA rejects this and counter-offers as in the previous case, and the game ends with the implied shares at $t = 1$.

The above result is very similar to the previous one in the sense that when the player is sophisticated, quasi hyperbolic discounting has no effect on the subgame perfect equilibrium except for a change in the discount rate. So, sophisticated players can be treated as exponential players with a lower discount factor. The second part of this proposition, again, follows from the lemma 1.

In this context, all the partially naive players’ offers will be rejected by EA and SHA. So, as long as $\hat{\beta} \neq \beta$, there will be one period delay. Among the partially naive types, the ones with higher $\hat{\beta}$ do better than the ones who have lower $\hat{\beta}$.

The results of the other cases where players are either exponential or sophisticated can be found easily by using the result in the remark. The NBI payoff figure 2.2 shows the payoffs of different types of agents in an alternating-offers bargaining game by using NBI as the solution concept. In Figure 2.2, Player 1 is the first proposer and Player 2 is the second proposer. The first expression in each box represents the share of Player 1 and the second expression represents the share of Player 2.

Now, comparison of the payoffs of each player is possible. These payoffs imply the following theorem:
48

**Theorem 1:** Assume each agent has the same time consistent impatience, δ. Further assume that if both agents have self-control problems, then their β’s are also the same. Then, Player 2’s (as an NHA) share is always greater than Player 2’s (as an SHA) share except when Player 1’s type is not NHA. Moreover, the players with higher \( \hat{\beta} \) get higher shares than the ones with lower \( \hat{\beta} \).

This theorem implies that an agent’s share as a second proposer is inversely related with awareness of self-control problems. This is valid for the first proposer case for some specific parameter values. Specifically, the more one agent is aware of self-control problems (the lower \( \hat{\beta} \in (\beta, 1) \)), the less share is obtained.

### 2.5 Equilibrium With Learning

Now the other two approaches, namely immediate and gradual learning, are considered. 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 (Lemma 1 actually extends this to other types.) With exponential and sophisticated agents, as the first proposer, the reason that for obtaining an
immediate agreement result without learning is the persistence of the NHA in being naive (persistent optimism about the self being time consistent in the future). No matter how the game proceeds, this optimism will persist. More importantly, since the exponential agent and sophisticated agent believe that the naive agent will stay naive whatever happens, they give up, in some sense, and offer a share that confirms the NHA’s incorrect self-belief.

Yildiz (2003) examines the effect of optimism about making offers in the future in a sequential bargaining model. He shows that if players will remain sufficiently optimistic for a sufficiently long future, then they will agree immediately in equilibrium. 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 is shown 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 unless the NHA is the first proposer. Since the naive player offers the opponent (either EA or SHA) less than what the opponent would accept, an offer that will be rejected, there is one period delay. Moreover, when both players are naive, they would never agree.

If the players are allowed to learn as they play the game and the EA holds optimistic beliefs about the naive player that the NHA may change self-beliefs, then it may be optimal for the EA to wait. The EA may wait for the NHA to be more sophisticated over time by making rejected offers since the EA gets more from a more sophisticated agent. Given the beliefs, the EA decides what to do based on the trade-off between the cost of waiting and getting a higher expected share. Thus, the EA will wait to persuade the NHA and the NHA will play the game as the EA wants it played.

As stated in O’Donoghue and Rabin (2001), delaying forever or sticking to the same belief about one’s own preferences 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, 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 self-beliefs. While generalization of this learning process is questionable in real-life, such learning can be obtained in specific
strategic environments (see the "health club" example in the introduction.)

Yildiz (2004) allows players to be optimistic about their bargaining power (measured as the probability of making offers), but they can also learn as they play the game. Here, the rational agent is allowed to hold optimistic beliefs about the NHA and the NHA is allowed to update self-beliefs over time. The NHA is not aware of the update in self-beliefs. In fact, this unawareness may lead to a delay in bargaining. Since the EA is optimistic about the NHA’s belief updating, the EA makes unexpected offers that will be rejected to force the NHA to update self-beliefs. A sequence of rejected offers will make the NHA realize her self-control problems, and cause the NHA to become more sophisticated.

2.5.1 Immediate Learning

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

Proposition 3: Let an EA and an NHA play the alternating-offers bargaining game. Assume the NHA immediately becomes sophisticated after seeing a rejection.

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 a 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. If the NHA offers first, then there will always be a one period delay.

Proof. 1. If the EA offers less than $1 - x = \frac{\beta \delta}{1 + \delta}$ to the NHA, then the NHA will reject it with the hope of getting at least this much share of the pie next period. If the NHA makes a rejected offer, then the NHA immediately becomes an SHA. Then, the NHA (now SHA) offers $y = 1 - \frac{1 - \delta}{1 - \beta \delta^2}$ 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 with a one period delay with $y = 1 - \frac{1 - \delta}{1 - \beta \delta^2}$. 
2. This part follows directly from the Lemma 1 where \( \hat{\beta}_{ji} = \beta_i < \hat{\beta}_i = 1 \) and \( \beta_j = \hat{\beta}_{ij} = \hat{\beta}_j = 1 \).

\[ \]

2.5.2 Gradual Learning

To make the learning process clear, the following examples can be given:

**Example 1:** Think about a student who decides to buy a pass for the school gym for regular attendance (to lose weight or for bodybuilding). This agent has optimistic belief in continuing to go to the gym regularly. However, it turns out that this is not possible for different reasons (other activities, boredom, laziness etc...) At the beginning of the next semester, the student will again consider buying the pass, but due to the past experience (in other words, realization of the self-control problem to some extent), different commitments are made (purchase with a friend so they can go together) or sufficient awareness occurs such that the pass is not purchased at all. However, the behavior pattern continues again and again, eventually buying gym pass ceases.

**Example 2:** Think about a student who registers for morning course sections. The student is optimistic about waking up early enough and attending classes regularly. However, if this is not possible for different reasons (other courses, laziness etc...), at the beginning of the next semester, the student will again consider registering morning sections, but due to the experience from the past, the student registers for fewer morning sections. However, if the same experience continues, again and again, eventually registering for any morning sections ceases.

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 and partially naive agents will be sophisticated over time in all situations that they face. However, in particular environments, naive agents may actually learn and update their self-beliefs. 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.

The assumption is that partially naive agent holds an initial belief (probability of using discount factor \( \beta \delta \) in the future) about the future self-control problem. This agent does not think that updating self-belief will occur, but when facing
rejections, updating occurs. The partially naive agent’s self-belief has a beta distribution that is widely used in statistical learning models.

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. The assumption is that the initial belief at $t = 0$ that impatience will occur at every future period is $\frac{m_{\beta \delta}}{n}$. After $m$ rejections, self-belief becomes $\frac{m_{\beta \delta} + m}{n + m}$ 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 $\alpha$ that measures the probability of the agent using $\beta \delta$ at any date $t$, and $\alpha$ is distributed with a beta distribution with hyperparameters $m_{\beta \delta}$ and $n$.

What is described in the last paragraph is not common knowledge. A partially naive agent is not aware of this updating structure but the rational agent knows this, and acquisition of this knowledge is the main 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:

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

This implies that the partially naive agent’s perception of $\beta$ is fixed and equal to $\hat{\beta}$ for the entire future. The intrinsic assumption here is that the partially naive agent is not perfectly forward looking. When this agent observes a rejection,

---

8 A model with more sophisticated naive players would be one we had thought about and one explicitly formulated by one of the referees for this paper.

9 Suppose that each player knows his type in the current period and has a conjecture about his future selves. Each future self of each player can be either exponential type or hyperbolic type. Also assume that the each player’s self-beliefs and about his opponent are common knowledge, hence the common-prior assumption is dropped. An exponential player is certain that each of his selves is exponential. A sophisticated player is certain that each of his selves is hyperbolic. A naive player is certain that each self is exponential. On the other hand, a naive learner assumes that a beta distribution with unknown mean generates the type of her future selves. Now, backwards induction or iterated conditional dominance can be applied and the analysis will be standard. A naive learner is now more “sophisticated” in the sense that she will be aware that she is learning overtime.

10 A reasonable expectation is that naive agent’s behavior becomes more like that of a sophisticated player. There might be a cutoff value of the parameters such that on one side the player would behave as if sophisticated and on the other side as if naive. The settlement date in this model might be earlier than the one in the original model. However, this formulation looks, at this point, to be analytically messy to solve and it is deferred for future research.
beliefs change, but no consideration is given to the possibility that future beliefs may change. Here, being partially naive 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). This specification also implies that the agent’s belief about date 50 at \( t = 49 \), which depends highly on the history, will be quite different from the belief about date 50 at \( t = 0 \). In addition, this learning specification implies that learning slows down over time (This means that the probability of being \( \beta \delta \) type is a function of \( m \), and the first derivative of this function with respect to \( m \) is positive, and second derivative is negative).

Note that perceived discount factor of the naive agent, \( \delta_t \), 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 more and more sophisticated (as \( m \to \infty \), \( \delta_m \to \beta \delta \)). Under this learning scheme, a person is said to be completely naive if \( 1 \leq m_{\beta \delta} << n \to \infty \). In other words, a larger \( n \) implies a more severe naivete. For the sake of the learning argument to be significant, the assumption is that initial belief of the agent is not so firm. In the case where the agent is too firm, the likelihood of obtaining a delayed agreement is low. A result that relates the firmness level with the extent of bargaining delay appears later.

Under these kinds 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 expected 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. By lemma 1, the partially naive agent’s offers are never accepted. This is because by rejecting the offer, the EA makes the NHA a little more sophisticated that allows the former to extract more from the latter in the next period.

The argument in the lemma 1 raises an immediate question as to why the naive agent cannot learn to make accepted offers. This is due to the following: The belief of the naive agent about the opponent’s belief, that the naive agent has a fixed
belief, does not change. The naive agent believes that the opponent holds beliefs that are the same as the naive agent’s self-beliefs. In other words, the naive agent infers nothing from the opponent’s offers about what the opponent thinks about the naive agent.

The partially naive player is not perfectly firm \((\frac{m\beta}{n} > 0)\). The agent updates self-belief about self-preferences in the future based on observed rejections, 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. The basic intuition behind this is that the exponential agent can extract more share from the partially naive agent by making her more sophisticated. This can be achieved by rejecting offers and delaying the game. However, since delay is costly, when the cost of delaying exceeds this benefit, the exponential agent finishes the game. Thus, this trade off between benefit and cost of delaying motivates the following theorem.

**Theorem 2:** 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.

In brief, Theorem 2 states that if the exponential agent holds optimistic beliefs about the partially naive agent, and if the partially naive agent learns over time, 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 or \(n\) is sufficiently high then, there will be an immediate agreement (e.g., \(t^* = k^* = 0\)).

**Corollary 2:** For any given \(m\beta\), \(\beta\) and \(\delta\), there exists some \(n^*\) such that \(\forall n \geq n^*\); the players reach an agreement immediately in the case that the exponential player is the first proposer and if the partially naive agent is offering first, then there will be one period delay by lemma 1.

**Proof.** Given any \(m\beta\), \(\beta\) and \(\delta\), in order for players to reach an immediate agreement, the following condition must be satisfied \(\forall k:\)

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

Define the following function:

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

where \(x_{2k}(\delta_{2k}) = 1 - \beta \delta (1 - \delta)\) and

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

Note that \(\forall k, \delta_{2k} \to \delta\) and \(\frac{x_{2k}(\delta_{2k})}{x_{2k+2}(\delta_{2k+2})} \to 1\) as \(n \to \infty\); also \(F(\delta; \beta, n, m_{\delta\beta}, k) < 0\) implies immediate agreement. As \(n \to \infty\), \(F(\delta; \beta, n, m_{\delta\beta}, k) \to \delta^2 - 1 < 0\). Thus, there exists an \(n^*\) such that for any \(n \geq n^*\), the above condition is satisfied. This is the case where the exponential agent is the first proposer. If the partially naive agent is the first proposer, then by Lemma 1, for any \(n \geq n^*\) the exponential agent rejects partially naive agent’s offer at \(t = 0\) and by the above argument, at \(t = 1\), players immediately agree.

Figure 2.3 shows delay depending on \(\delta\) and some specific parameter values \(\beta = 0.5, m_{\delta\beta} = 1\) and \(k = 0\). As the corollary suggests, for any given patience level, \(\delta\), a firmness level can be found, \(n\), such that players immediately agree in the bargaining game \((F(\delta; \beta, n, m_{\delta\beta}, k) \leq 0, e.g., \text{given } \beta = 0.5, m_{\delta\beta} = 1\) and \(\delta = 0.995, \text{for all values of } n \text{ satisfying } n \geq n^* = 3050, \text{immediate agreement occurs.})\)

On the other hand, the following corollary shows that as long as the exponential agent is patient enough, there will be delay.

Corollary 3: For any given \(m_{\delta\beta}, \beta \in (0, 1)\) and \(n\) satisfying condition 1 below, there exists a \(\delta^*\) such that for every \(\delta > \delta^*\), there will be \(2k + 2, k = 0, 1, 2, \ldots\), period delay (or \(m = 2k + 2\) rejections) in the case that the exponential 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.

\[\beta \geq \beta^* \text{ where } \beta^* = \frac{(m_{\delta\beta} + 2k)(m_{\delta\beta} + 2k + 2)}{(m_{\delta\beta} + 2k)(m_{\delta\beta} + 2k + 2) + (n - m_{\delta\beta})} \quad (\text{Condition 1})\]
Figure 2.4 shows delay depending on \( n \) and some specific parameter values \( \beta = 0.5, m_{\beta \delta} = 1 \) and \( k = 0 \). As this corollary suggests, for any \( n \), a patience level \( \delta \) can be found such that delay occurs, \( F(\delta; \beta, n, m_{\beta \delta}, k) > 0 \). For example, given \( \beta = 0.5, m_{\beta \delta} = 1 \) and \( n = 100 \), all values of \( \delta \) satisfying \( \delta > \delta^\ast = 0.9778 \), \( k = 0 \) produces, at least, a 2 period delay. \( \beta^\ast = 0.6 \) for \( n = 3 \). Since, in the example here, \( \beta = 0.5 < 0.6 = \beta^\ast \), for \( n = 3 \), there does not exist any \( \delta^\ast \) such that for every \( \delta > \delta^\ast \), \( F(\delta; \beta, n, m_{\beta \delta}, k) > 0 \). Thus, for \( \beta = 0.5, m_{\beta \delta} = 1, n = 3 \) and for every \( \delta \in (0, 1) \), there is immediate agreement.

### 2.6 Conclusion

The literature on non-cooperative games played by time inconsistent agents and learning of naive time inconsistent agents over time is very limited. To fill this gap, this study presents the consequences of incorporating time inconsistency and naive learning into a sequential bargaining model where some players are time inconsistent and are able to learn their preferences over time. This paper models time inconsistency by using quasi hyperbolic discounting.

Under the complete information assumption, I first show that offers of the players who are naive or partially naive are never accepted by any type of player in both no learning and gradual learning cases. Then, by following the traditional approach that considers naive agents who never learn their types, I show that when a naive agent makes the first offer to an exponential or a sophisticated agent, one period delay occurs. In general, the payoff of the partially naive agent increases with the degree of naivete. Moreover, without learning, two naive agents disagree forever. The rationale behind this result is the persistency of the naive agents in being optimistic about their future self-control problems.

On the other hand, if the naive agent is able to learn to be more sophisticated over time, it is possible to obtain multiple periods of delay in bargaining played between the naive agent and the exponential agent. Given that the exponential agent gets a higher share from a more sophisticated agent, the exponential agent has an incentive to make the naive agent more sophisticated by delaying the game. However, delaying is costly. The trade-off between the cost of delaying and the incentive to delay characterizes the optimal delay time from the perspective of the
exponential agent. When the partially naive agent is close to being completely naive, there will be an immediate agreement. Conversely, if the exponential agent is patient enough, then the agreement will be delayed.

Exploring the behavioral characterization of time inconsistent agents in more general settings still remains as an issue that needs further research. More specifically, it would be interesting to explore the issue of incomplete information about the opponent’s type and reputation formation. Additionally, the learning dynamics and information processing structure can be relaxed to obtain more general results.

APPENDIX

Proof of Lemma 1: Before starting the proof, it is important to mention the assumption, again, that each naive and partially agent recognizes the other player’s time inconsistency and potential learning ability completely but not fully that of the self.

There are two players $i$ (she) and $j$ (he.) Player $i$ makes the first offer and has $\beta_i \in (0, 1)$, $\delta \in (0, 1)$ and $\hat{\beta}_i \in (\beta, 1)$. This means that player $i$ has time inconsistency problem, and she $i$ is naive or partially naive. Player $j$ has $\beta_j \in (0, 1]$, $\delta \in (0, 1)$ and $\hat{\beta}_j \in [\beta, 1]$. This means that player $j$ may have time inconsistency problem and if he has, he is sophisticated, naive or partially naive.

To incorporate immediate and gradual learning in the analysis, I define the following: $\hat{\beta}_{ij}$ represents the belief of agent $i$ about what $j$’s perception ($\hat{\beta}_j$) will actually be if there is a rejection\(^{11}\). $\hat{\beta}_{ji}$ is defined similarly. Note that $\beta_i \leq \hat{\beta}_{ji} \leq \hat{\beta}_i$ with $\beta_i < \hat{\beta}_i$ and $\beta_j \leq \hat{\beta}_{ij} \leq \hat{\beta}_j$. Without learning, $\hat{\beta}_{ji} = \hat{\beta}_i$ and $\hat{\beta}_{ij} = \hat{\beta}_j$. The conditions $\beta_{ji} \leq \hat{\beta}_i$ and $\beta_{ij} \leq \hat{\beta}_j$ are satisfied with learning because each agent knows that the opponent (naive or partially naive) is learning so the opponent’s perception will be closer to what it actually is, which is smaller than the today’s perception.

What follows is an examination of what each agent offers. At $t = 0$, agent $i$

\(^{11}\)The actual value of $\hat{\beta}_{ij}$ depends on the number of rejections. However, since one’s decision today, among all the future possible perceptions of the opponent, depends only on the opponent’s perception tomorrow, defining $\beta_{ij}$ like this is enough for the analysis.
offers \((x, 1 - x)\), where \(x\) is \(i\)'s share. \(i\) will make this offer such that

\[
1 - x = \beta_i \delta (1 - y)
\]

where \((y, 1 - y)\) is the offer of \(j\) at \(t = 1\), where \(y\) is \(i\)'s share. At \(t = 0\), agent \(i\) thinks that \(j\) will get at \(t = 1\), \(1 - y = \frac{1 - \hat{\beta}_j \delta}{1 - \beta_i \hat{\beta}_j \delta^2}\) based on the belief that \(j\) will have \(\hat{\beta}_j\) after tomorrow and \(i\) will have \(\hat{\beta}_i\) after tomorrow. Thus, \(i\) offers \(j\)

\[
1 - x = \beta_j \delta \left( \frac{1 - \hat{\beta}_j \delta}{1 - \beta_i \hat{\beta}_j \delta^2} \right)
\]

On the other hand, \(j\) knows that payoffs at \(t = 1\) and at \(t = 2\) will be discounted with \(\beta_i \delta\) by \(i\). He also knows that \(i\)'s perception tomorrow will be \(\hat{\beta}_j\). Then, at \(t = 1\), he expects to get

\[
1 - y = 1 - \beta_i \delta \left( \frac{1 - \hat{\beta}_j \delta}{1 - \beta_i \hat{\beta}_j \delta^2} \right)
\]

Now, agent \(j\) compares the offer today, \(1 - x = \beta_j \delta \left( \frac{1 - \hat{\beta}_j \delta}{1 - \beta_i \hat{\beta}_j \delta^2} \right)\) and what \(j\) expects tomorrow, \(1 - y = 1 - \beta_i \delta \left( \frac{1 - \hat{\beta}_j \delta}{1 - \beta_i \hat{\beta}_j \delta^2} \right)\).

\[
1 - x = \beta_j \delta \left( \frac{1 - \hat{\beta}_j \delta}{1 - \beta_i \hat{\beta}_j \delta^2} \right) \text{ today and } 1 - y = \beta_i \delta \left( \frac{1 - \hat{\beta}_j \delta}{1 - \beta_i \hat{\beta}_j \delta^2} \right) \text{ tomorrow} \Rightarrow
\]

\[
\beta_j \delta \left( \frac{1 - \hat{\beta}_j \delta}{1 - \beta_i \hat{\beta}_j \delta^2} \right) \text{ and } \beta_i \delta \left( \frac{1 - \hat{\beta}_j \delta}{1 - \beta_i \hat{\beta}_j \delta^2} \right) \Rightarrow
\]

\[
(1 - \hat{\beta}_j \delta)(1 - \hat{\beta}_j \delta^2) \text{ and } (1 - \hat{\beta}_i \hat{\beta}_j \delta^2)(1 - \hat{\beta}_j \delta^2 - \beta_i \delta + \hat{\beta}_i \hat{\beta}_j \delta^2)
\]

Rearranging the terms will give the following

\[
\hat{\beta}_j \delta^2 (\hat{\beta}_i - \beta_i) + \hat{\beta}_i \hat{\beta}_j \delta^3 (\hat{\beta}_ji - \beta_i) \text{ and } \delta (\hat{\beta}_i - \beta_i) + \hat{\beta}_i (\hat{\beta}_j) \delta^4 (\hat{\beta}_ji - \beta_i) \Rightarrow
\]

\[
\hat{\beta}_i \hat{\beta}_j \delta^2 \text{ and } \frac{(\hat{\beta}_i - \beta_i)}{\hat{\beta}_ji - \beta_i}
\]

Since \(1 \leq \frac{(\hat{\beta}_i - \beta_i)}{(\hat{\beta}_ji - \beta_i)}\) is always satisfied and \(\hat{\beta}_i \hat{\beta}_j \delta^2 < 1\), \(\hat{\beta}_i \hat{\beta}_j \delta^2 < 1 \leq \frac{(\hat{\beta}_i - \beta_i)}{(\hat{\beta}_ji - \beta_i)}\).
Thus, rejecting today is an optimal strategy.

Each case is specified as follows:

**Without Learning:**

Note that, without learning, $\hat{\beta}_{ji} = \hat{\beta}_i$ and $\hat{\beta}_{ij} = \hat{\beta}_j$. Player $i$ is naive or partially naive, $\hat{\beta}_i \in (\beta_i, 1]$, and $j$ can be either one of the four types. At each case, the analysis above is valid and gives at least one period delay result. Except in the case where $\hat{\beta}_j = \beta_j \neq 1$ or $\hat{\beta}_j = \beta_j = 1$, the agents never agree.

**Immediate Learning:**

With immediate learning, $\hat{\beta}_{ji} = \beta_i$ for every type $j$ and $\hat{\beta}_{ij} = \beta_j$. The proof applies to these cases too with the above specifications. After one rejection, agent $i$ becomes sophisticated ($j$ also becomes sophisticated if $\hat{\beta}_j \in (\beta, 1]$.) Then, at $t = 1$, both are sophisticated and they agree immediately. If $j$ is already sophisticated or exponential at $t = 0$, then they agree at $t = 1$ (see proposition 3, part 1.)

**Gradual Learning:**

With gradual learning, both $\hat{\beta}_{ji}$ and $\hat{\beta}_{ij}$ are functions of the number of rejections and they are decreasing with it. As time passes $\hat{\beta}_i \to \beta_i$ and $\hat{\beta}_j \to \beta_j$. Also, at each period $\beta_i < \hat{\beta}_{ji} < \hat{\beta}_i$ and $\beta_j < \hat{\beta}_{ij} < \hat{\beta}_j$. For the cases where $\hat{\beta}_j = \beta_j \neq 1$ or $\hat{\beta}_j = \beta_j = 1$, there will be finite periods of delay (no delay is possible. For this argument, see theorem 2.) For all the other cases, there will be no agreement because of the followings: 1. Neither agent will be fully sophisticated as a result of the learning process. 2. Even without learning, each naive or partially naive agent’s offer is rejected by any type of the opponent. With learning where each player has relatively higher incentive to make rejected offers and to reject even lower offers than what is expected, it would be easier to get perpetual disagreement.

The detailed proof of the lemma (for a game played by a partially naive agent and an exponential agent) with specific values for $\hat{\beta}_{ji}$‘s is added here to make the above argument clearer.

At any time $t = 2k$, $k = 0, 1, 2...$, the partially naive agent ($i$) makes an offer. The EA ($j$) is the second proposer. Since the $\hat{\beta}_{ji}$‘s depend on the number of rejections (so, on the time period), I call the partially naive agent’s offers as $x_{2k}(\delta_{2k})$ at $t = 2k$ and the EA’s offers at odd periods as $y_{2k+1}(\delta_{2k+1})$ where $x_{2k}(\delta_{2k})$ and $y_{2k+1}(\delta_{2k+1})$ are the partially naive agent’s shares. Reaching time $t = 2k$ means $m = 2k$ rejections occurred as of time $t$. Agent $i$ has been updating self-
beliefs during this time and now, i's perception on discount factors for any time $t > m = 2k$ is $\delta_{2k} = [1 - \frac{m\beta + 2k}{n + 2k} (1 - \beta)] \delta$ (for the details of the learning framework, see Section 5.2.) Note that $\hat{\beta}_{ij} = \hat{\beta}_i = \beta_j = 1$. At $t = 2k$, player $i$ discounts payoffs at $t = 2k + 1$ by $\beta\delta$ but considers discounting payoffs at each $t > 2k + 1$ by $\delta_{2k}$. On the other hand, $\hat{\beta}_{ji} = \delta_{2k+1} = [1 - \frac{m\beta + 2k + 1}{n + 2k + 1} (1 - \beta)] \delta$.

From previous arguments, $i$ offers the following to $j$ at even periods:

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

It is also known that at every odd period, $j$ will never offer more than $y_{2k+1}(\delta_{2k+1})$ to $i$ where

$$y_{2k+1}(\delta_{2k+1}) = \beta_i \delta \left( \frac{1 - \hat{\beta}_j \delta}{1 - \hat{\beta}_{ji} \hat{\beta}_j \delta^2} \right) = \beta \delta \left( \frac{1 - \delta}{1 - \delta \delta_{2k+1}} \right)$$

Now, it is enough to show that when $i$ offers, the discounted value of the minimal share in the next period that is acceptable by $j$ is greater than what $i$ offers to $j$ today, which is $\delta(1 - y_{2k+1}(\delta_{2k+1})) > 1 - x_{2k}(\delta_{2k})$.

$$\delta(1 - y_{2k+1}(\delta_{2k+1})) > 1 - x_{2k}(\delta_{2k})$$

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

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

Now, it is enough to show that the first inequality is satisfied because the second inequality is already satisfied since $\delta_{2k+1} < \delta_{2k} \forall k$.

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

$$\Rightarrow 1 - \delta \delta_{2k+1} > 1 - \delta_{2k+1} + \beta \delta - \beta \delta^2$$

$$\Rightarrow \delta_{2k+1} > \beta \delta$$

The last inequality is true by definition, so this implies $\delta(1 - y_{2k+1}(\delta_{2k+1})) > 1 - x_{2k}(\delta_{2k}) \forall k$. Thus, player $j$ rejects $1 - x_{2k}(\delta_{2k}) \forall k$. ■
Proof of Proposition 2:

1. When the SHA offers first, the offer to the NHA is the discounted value of what she thinks that she can get tomorrow. If the SHA offers less, then the NHA will reject this offer. To convince the NHA, he offers $1 - x = \beta \delta (1 - y)$ where $(1 - y) = \frac{1 - \beta \delta}{1 - \beta \delta^2}$ (this is the highest share that the NHA can get by rejecting today [check the first remark].)

2. This result again follows from the Lemma 1 where $\beta_i < \hat{\beta}_{ji} = \hat{\beta}_i = 1$ and $\beta_j = \hat{\beta}_{ij} = \hat{\beta}_j < 1$. ■

Proof of Theorem 1: If a player makes the second offer, then in terms of payoffs, if that player is EA, then she does better than the case where she is an NHA who does better than the case where she is an SHA (except when the opponent is also naive in which case they both end up with nothing). This is true since:

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

When the naive agent is the first proposer there is always delay. When the opponent is EA or SHA, there is only one period delay but when the opponent is also naive, there is no agreement. The theorem’s statement is valid for the first player also under some specific parameter values. Player 1’s share as an NHA is always greater than player 1’s share as an SHA except if Player 2’s type is not NHA, if

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

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 case too.

The more naive the player is, the higher share is received. The effect of naive is as follows: The partially naive agent’s perception of her self-control problems is $\hat{\beta}$. If naive agent is replaced with a partially naive agent in proposition 2 and 3, by using the same argument in proposition 2 and 3, the following results arise under the assumption that each agent knows the opponent’s type:
1. If the EA is the first proposer, then equilibrium shares are

\[(x^*, 1 - x^*) = (1 - \beta \delta \frac{1 - \delta}{1 - \beta \delta^2}, \beta \delta \frac{1 - \delta}{1 - \beta \delta^2})\]

where \(x^*\) is the share of the EA,

2. 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.

Comparing the shares of sophisticated, partially naive and naive agents reveals that the assertion above is true. Moreover, in both cases, if the partially naive agent is the first proposer, then there will always be one period delay for any value of \(\hat{\beta} \in (\beta, 1)\) by a similar arguments in propositions. ■

Proof of Theorem 2: By lemma 1, the partially naive agent’s offers will be rejected at each odd period. Then, the rational agent will compare following payoffs that can be obtained at each even period \((x_{2k}(\delta_{2k}), \text{ where } k = 0, 1, 2..., \text{ which is defined in the proof of Corollary 2)}:\)

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

Since the EA is time consistent and will make the same comparison at any given period of time, the EA’s choice will be the largest element of the above sequence today and the EA will implement what is decided today. The payoff that can be received is increasing in time but waiting is costly. Hence, there should be an optimal waiting time that allows extraction of the highest share 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, the EA offers \((x_{t^*}(\delta_{t^*}), 1 - x_{t^*}(\delta_{t^*}))\), the partially naive agent accepts and the game ends at \(t = t^*\). This threshold \(t^*\) exists because as time passes, learning slows down and the additional payoff that the EA expects by waiting because increased sophistication...
of the partially naive agent is offset by the loss of waiting. In any case of multiplicity of optimal value of $k$, minimum of those $k$ values is taken as $k^*$. 

The proof of the second part of the theorem (also the existence of a maximum for the first part) is as follows: Since both $\delta^{2k}$ and $x_{2k}(\delta_{2k})$ are strictly positive for all parameter values (except when $\delta = 0$, but I excluded this case), $\delta^{2k}x_{2k}(\delta_{2k})$ is also strictly positive. Let $f = \delta^{2k}$, $g = x_{2k}(\delta_{2k})$ and $h = f \ast g = \delta^{2k}x_{2k}(\delta_{2k})$. Then, at every $t > t^*$ players immediately agree, if function $h$ is a decreasing function of $k$ after it reaches its maximum (at some $k^*$).\footnote{It does not matter whether maximum occurs at an integer value of $k$ or not. We solve the problem for the real valued $k$’s and then take the nearest integer smaller than the $k^*$ value.} In other words, if $\forall k > k^*$, $\frac{d(h(k))}{dk} < 0$, then at every $t > t^* = 2k^*$, players immediately agree. This is enough to prove the second part of the theorem because if this is the case, then the EA does not have an incentive to delay the game anymore after $t > t^* = 2k^*$ since the EA’s expected discounted share (the value of function $h$) is decreasing. To show this, we will look at the rate of change of the functions whose multiplication gives the function $h$. The rate of change of any function $F$ is given as $F'$ where $F'$ is the derivative of the function $F$. So, the rate of change of function $f$ and $g$ are $\frac{df}{dk}$ and $\frac{dg}{dk}$, respectively where $f' = \frac{df}{dk}$ and $g' = \frac{dg}{dk}$. Note that $\frac{df}{dk}$ is a strictly negative constant, $\frac{df}{dk} = 2\ln(\delta)$. Also, $\frac{g'}{g}$ is always positive and converges to zero as $k \rightarrow \infty$. Then, $\frac{h'}{h} = \frac{(fg)'}{fg} = \frac{f'g + fg'}{fg} = \frac{f'}{f} + \frac{g'}{g}$. Since $\frac{f'}{f}$ is a strictly negative constant and $\frac{g'}{g} \rightarrow 0$ as $k \rightarrow \infty$, there exists a $k^*$ such that $\forall k \geq k^*$, $\frac{h'}{h} \leq 0$. This completes the proof. \[\blacksquare\]

**Proof of Corollary 3:** Given any $m_{\beta\delta}$, $\beta$ and $n$, there will be $2k + 2$ periods delay if:

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

Define function $F(\delta; \beta, n, m_{\beta\delta})$ as in the proof of Corollary 2:

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

(1)

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

$$+ \frac{\beta\delta(1 - \delta)}{1 - \delta^2[1 - (1 - \beta)(\frac{m_{\beta\delta} + 2k}{n + 2k})]}$$
Then, (1) implies that for any given \( \beta, m_{\beta \delta}, n \), if there exists a \( \delta^* \) such that for any \( \delta > \delta^* \)

\[
F(\delta; \beta, n, m_{\beta \delta}, k) > 0
\]

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 \( \delta \) 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 (2) is true. In other words, when \( \delta \) 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 \( \delta \) 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.

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

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

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

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

\[
\Rightarrow \beta > \frac{(m_{\beta \delta} + 2k)(m_{\beta \delta} + 2k + 2)}{(m_{\beta \delta} + 2k)(m_{\beta \delta} + 2k + 2) + (n - m_{\beta \delta})} \quad \text{(Condition 1)}
\]

Thus, for any given value of \( \beta \in (0, 1), m_{\beta \delta} \) and \( n \) satisfying Condition 1, there exists a \( \delta^* \) such that for any \( \delta > \delta^* \), there will be \( 2k + 2 \) periods delay in the bargaining game between the partially naive agent and the exponential agent when the exponential agent is offering first. When the partially naive player offers first, by lemma 1, there will be one more period of delay, \( 2k + 3 \) periods.
Figure 2.4 shows delay depending on $n$ and some specific parameter values $\beta = 0.5$, $m_{\beta \delta} = 1$ and $k = 0$. Since $\beta = 0.5 < 0.6 = \beta^*$, for $n = 3$, in the figure, there is no $\delta^*$ such that for every $\delta > \delta^*$, $F(\delta; \beta, n, m_{\beta \delta}, k) > 0$ and no delay occurs. For all $n$ such that $\beta > \frac{(m_{\beta \delta} + 2k)(m_{\beta \delta} + 2k + 2)}{(m_{\beta \delta} + 2k)(m_{\beta \delta} + 2k + 2) + (n - m_{\beta \delta})} \Rightarrow 0.5 > \frac{3}{2 + n} \Rightarrow n > 4$, there will be delay. Figure 2 gives specific values for $\delta^*$ for each $n$.

$\beta = 1$ case that satisfies condition 1 is excluded. However, when $\beta = 1$, $F < 0$ is always satisfied $\forall \delta(0, 1)$. This case is an exception because when $\beta = 1$, $\delta^* = 1$ and there is no $\delta$ satisfying $\delta > \delta^*$. In this case, two agents are exponential having $\delta$ time consistent discount factor. So, there will be no delay and the pie is shared as in the standard case.

For the case where $\delta = 1$, $F = 0 \forall \beta(0, 1)$. So, there will be no delay. The same is true for the case where $\delta = \beta = 1$, $F = 0$, too, resulting in no delay. ■

2.7 References


Figure 2.3. Delay depending on delta, beta=0.5, m=1, k=0

Figure 2.4. Delay depending on n, beta=0.5, m=1, k=0
3.1 Introduction

3.1.1 The Problem

There is a growing literature documenting the tendency of people to procrastinate when they face intertemporal decision problems. This kind of behavior creates some time inconsistency that cannot be completely explained by conventional economic theory. These agents are time-inconsistent in the sense that they are very keen on their near future selves’ gratification such that their distant-future selves will probably be regretful about this pursuit of immediate gratification. Due to this tendency, they behave differently than the way they planned to behave in the future.

The most obvious symptom of time inconsistent preferences is procrastination of unpleasant and costly tasks with the hope of completing them in the future. This procrastination tendency sometimes results in inefficient behavior such that even if taking an action is optimal, agent may procrastinate doing it. Writing a proposal for your thesis, finishing your paper, filing your taxes, starting a diet, attending gym regularly are some of the examples of costly tasks that we always have the tendency to delay up to the deadline (handing in your proposal at the very last
moment, long lines in front of post office on April 15, etc.) or to procrastinate in a nonreversible way (not quitting smoking and die early, not going gym regularly and get fat, not filing your taxes and lose the tax return, etc.)

Although inefficient procrastinative behavior is observed, there is still controversy in explaining it. Mukherji et al [2002], Rubinstein [2003], Dasgupta and Maskin [2002] all give alternative explanations why such behavior might be observed. A different point of view is taken by Laibson [1997] and by O’Donoghue and Rabin [1999, 2003]. We adopt their formulation of time inconsistent preferences, which they call quasi hyperbolic discounting. We do not, in this paper, express any opinion about the foundational questions involved; the Laibson-O’Donoghue-Rabin formulation poses some interesting questions that we seek to answer.

The problem we are going to consider consists of two stages. In the first stage, an individual agent, whose time preferences exhibit quasi hyperbolic discounting, has to choose when to complete a sequence of costly, inalienable investments in his human capital. After completing all the investments, a surplus is generated for the employer, if the agent uses his newly acquired skills. The employer and the agent bargain over the division of this surplus at the time the sequence of investments is completed. Thus, this time is endogenous in our model. Also, the employer knows the nature of the agent’s time preferences. Bargaining takes place through a Rubinstein’s alternating offers procedure with the discounting being standard exponential for the employer and quasi hyperbolic for the agent. The introduction of two players in this setting that includes bargaining is new in this paper so we refer the reader to the previous chapters for the discussion of the equilibrium concept. We work backwards to solve the game. We first examine the bargaining stage and then determine the behavior of the quasi hyperbolic agent in the investment stage who can be either naive or sophisticated. We do not, in this paper, consider partially naive agents as in O’Donoghue and Rabin do [2001].

3.1.2 Examples

Example 1: Think about a doctor or a technician who wants to work in a hospital. There are some requirements of the hospital that the doctor must fulfill. She must pass some eligibility exams, have some kind of special training and have
some certificates before she goes to interview with the hospital. These activities that she must complete are costly tasks including self investments. If the agent (doctor) has time inconsistent preferences, then how is she going to make these investments? Is she going to finish all tasks in an efficient way or is she going to have some procrastination motive that makes her postpone the investments?

Example 2: Think about a junior employee in an insurance firm who has to complete several actuarial examinations before she can undertake some assignments for the company. In our model, the employee has time inconsistency problems but the employer is time consistent. Preparing for the exams is a costly task that needs some work such as reading and research. In this situation, how does she allocate her time on these costly tasks? Does she procrastinate taking the exams or does she invest herself quickly and take the exams as soon as possible?

Example 3: Think about a university student. She has the opportunity to take extra classes such as computer programming, leadership, management outside her major. By taking these courses, she increases her chance to get paying internships in summers. Moreover, if she works in summers, then she can get a job more easily in the companies she interned in or in other companies after graduation. However, taking those courses is costly in terms of time, effort and money. If she has time inconsistent preferences, does she take those classes at each semester regularly? does she postpone taking them? Or does she not take them at all?

People’s lives are full of the problems like ones above in which they always face the trade-off of finishing long term projects and their immediate costs. In this paper, we will try to explain the people’s sometimes inefficient behavior in these kinds of situations.

3.1.3 Related Literature

First, Strotz [1956] suggested that people are more impatient when they make short-run trade-offs than when they make long-run trade-offs. When two payoffs are both far away in time, the decision maker tends to be more patient, on the other hand, when two payoffs are relatively close in time, then the decision maker is likely to behave more impatiently (e.g., choosing between “15 minutes break now, today or 1 hour break tomorrow” and “15 minutes break in 30 days or 1 hour break
in 31 days”, in this example, people -of course, generally- prefer the first choice in the first offer and second in the second offer which reflects the idea of discounting future in a different way). Briefly, we wish to act patiently in the long run but the desire for instant satisfaction frequently overwhelms our good intentions.

Traditionally, it is assumed that discount factors are exponential that means a util delayed $\tau$ periods is worth $\delta^\tau$ as much a util enjoyed immediately ($\tau = 0$). Some of the hyperbolic discount functions that are used in the literature, Chung (1961), are $1/\tau$ and $1/(1 + \alpha \tau)$ with $\alpha > 0$. Laibson [1997a] used called “quasi hyperbolic discount” function, $\{1, \beta \delta, \beta^2 \delta^2, \beta^3 \delta^3, \ldots\}$ where $\beta \in [0, 1]$ and $\delta \in [0, 1]$.

Since this is a relatively new approach, there is controversy about whether we really observe hyperbolic discounting in the consumer behavior, Dasgupta and Maskin [2002]. Because this is not a foundation, this can only be argued via experiments on people and by some real life facts. There is a growing literature in this aspect of the issue both in favor of and against it too, see Mukherji et al [2002] and Rubinstein [2003] for a more detailed discussion.

The second stage game is a bargaining game between a time consistent principal and a time inconsistent hyperbolic agent. We apply the alternating offer bargaining framework proposed first by Rubinstein [1982]. Rubinstein assumes stationary preferences over time. 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. They find that when the time interval between offers goes to zero, there exists a unique subgame perfect equilibrium.

Behavioral characterization of economic agents in the context of costly investment under the assumption of time inconsistent preferences was examined by O’donoghue and Rabin in a series of papers, [1999a], [1999b], [2001], [2003]. In their paper called “Doing It Now or Later” [1999a], they assume two different characteristics of time inconsistent behavior, naive and sophisticated, and two different cost and reward structures, immediate rewards and immediate costs. Naive hyperbolic agent, NHA, is not aware of her future preference reversals or self control problems, however, sophisticated hyperbolic agent, SHA, is fully aware of her self control problems so that she predicts how her future selves will behave in the future correctly. O’Donoghue and Rabin find that naive agents procrastinate
immediate cost activities and do immediate reward activities too soon. Whereas, Sophistication relieves procrastination and exacerbates preproperation (in their terminology). In addition, when there are multiple tasks, there is no general result saying that sophisticates always finish the tasks before naives as in the case of only one task.

In the next paper [1999b], they introduce a model in which principals can design incentives to induce time inconsistent procrastinators to complete tasks in an efficient way. Risk neutral agents face a task having a stochastic cost structure. They find that if task-cost distribution is common knowledge, the efficient outcome, which minimizes the sum of the task cost for agent and waiting cost for principal, can be achieved. If task-cost distribution is only known by the agent, efficiency often cannot be achieved for procrastinators. Also, they show that optimal incentives for procrastinators involve an increasing punishment for delay as time passes.

Partial naivete (neither completely naive nor completely sophisticated) and menu of tasks for economic agents are introduced in their “Choice and Procrastination” paper, [2001]. They basically show that providing a nonprocrastinator additional options can induce procrastination and a person may procrastinate worse while pursuing important goals than unimportant ones.

When time inconsistent agents face long term projects rather than projects that are completed once begun, they can choose when and whether to complete each stage of the projects. In their recent paper, [2003], O’Donoghue and Rabin show that not only procrastination in starting the project but also never completing the project, even if some cost was already incurred, can be observed for naive agents. Cost distribution is the key in determining the behavior of agents in this environment. If the cost structure is endogenous, then people are more likely to choose cost structures that make them start but not finish the project.

3.1.4 Contributions of This Paper

We will use Laibson’s quasi hyperbolic discount function in this paper. Since Laibson’s representation is widely used in modeling time inconsistent behavior, we will take this as given in our paper and we will apply it to our framework without
arguing whether it is the true approach or not. In O’Donoghue and Rabin’s 2003 paper, there are two stages of investment. While we also have the sequence of investments (here we generalized it to k units) they have, in their case the wage after completion is exogenously given. However, the preferences and types also affect the payoff. In our model, the payoff is made endogenous by introducing bargaining as the second stage game. We believe that this approach carries O’Donoghue and Rabin’s framework to a broader context and it is more realistic and plausible in such situations that have not only interactions among the agents’ selves but also strategic interaction of different agents. Introducing the second stage bargaining game, as well as introducing boundedly rational players in a bargaining game are new approaches. So, one of the most interesting aspects of our problem is the interplay between alternating offers bargaining and time inconsistent preferences. Therefore, our version of subgame perfect equilibrium is also somewhat new. This is a twist from the classical rational expectations approach because the beliefs of time inconsistent agent turn out to be wrong eventually.

We do not give a formal definition of the equilibrium and will refer to it as subgame perfect equilibrium in the sequel, the definition of which is well known. Rational expectations that is used in almost all strategic games assumes that people do not make systematic errors when predicting the future and forming beliefs about the future. Obviously, Naive hyperbolic agent (NHA) is systematically wrong about her behavior in the future and also about the opponent’s belief about her. So, they are boundedly rational. However, the crucial thing in this strategic environment is mutually giving best responses even if there exists this kind of inconsistent beliefs. 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 behavior, the last requirement is difficult to satisfy. In our case, the beliefs about the other player’s actions are correct but the hyperbolic discounter believes that the exponential player has wrong beliefs. (Since the time inconsistent player is wrong about his own future actions, it is not surprising that he believes the exponential player is also wrong about these.) Given this caveat, subgame perfectness is defined in the usual way as being Nash after every history.

In this nonstationary environment, since beliefs about the actions and the ac-
tual actions may differ, it is crucial to look at the beliefs of each agent. In the bargaining game, the beliefs are like the following: Exponential agent, EA, believes that "NHA believes that I am EA" and "NHA will have this self-control problem not only in the very near future but also in the distant future" and she is right in these beliefs. On the other hand, NHA believes that "the opponent is actually EA", "Today, I follow my immediate gratification but this will not be the case in the future" and "EA believes that I am naive and I will behave time inconsistently in the future but he is wrong about this". However, NHA is wrong in all these beliefs except the opponent’s type. She thinks that EA is wrong about herself but actually EA is right. She is wrong about herself too.

Again, the equilibrium concept includes the mutual best responses given the above beliefs. First, they do not want any delay in agreement. Second, since in equilibrium the game will end in the first period, the wrong beliefs of NHA will not be implemented. However, the equilibrium shares are determined based on these beliefs. If EA makes the first offer, then he will offer NHA a share that will be accepted given NHA’s wrong beliefs about herself. NHA will accept this offer because she thinks that she will be time consistent in the future and EA takes this into account while making the offer.

In our paper, we have an exogenous time constraint for the completion of the project. As an extension, O’Donoghue and Rabin think a partial reward scheme (payoff is given not only when the investment stage is finished but also after each unit is invested) under the assumption of fixed total reward (partial rewards are basically transfers from the total payoff in the end) that causes more severe procrastination. However, we will consider a variable (not constant) total reward scheme as an incentive mechanism to make agents not to procrastinate by introducing some ”bonus” scheme for each invested unit (the principal actually gives up from some of her surplus that she will earn in the end).

One other assumption is that the type of each agent is released in the bargaining stage. The rationale behind this is that in the first stage, the investment pattern of each agent signals the type of the agent, so in bargaining game that is played right after the investment stage is finished, there is complete information about the types of the agents.

The other interesting point of introducing the bargaining stage is that since
NHA is mistaken in predicting her wage (she predicts a higher wage than she will actually get, which shall be clear later) as the result of second stage game, this may create a motive of regret that she may finish a project that is not worth finishing since she will get a smaller payoff than she expects. So, this may be an interesting approach in understanding people’s disappointments resulting from their great expectations about the future.

Moreover, we assume an infinite and discrete time horizon but we assume an exogenous deadline for the first stage of the game rather than assuming a limitless investment phase. However, this assumption is plausible in some situations but it is not in others. So, it can be relaxed according to the characteristic of the interested problem. In this long term project framework, O’Donoghue and Rabin also examine the behavioral consequences of the partial naivete, which is defined and explained in [2001], and find that degree of naivete alters the investment patterns of the agents. We can easily extend our analysis to that case too.

3.1.5 Description of the Model and Preliminary Results

We will give an informal description of the model and state preliminary results that we get. In order to make the comparison between different characteristics of preferences in our model, we will suppose that the principal is an exponential discounter and the agent is either exponential agent (EA, time consistent), naive hyperbolic agent (NHA) or sophisticated hyperbolic agent (SHA). Exponential discounter has a sequence of discount factors \( \left\{ 1, \delta, \delta^2, \delta^3, \ldots \right\} \). Naive and sophisticated hyperbolic agents have the same sequence of discount factors \( \left\{ 1, \beta\delta, \beta\delta^2, \beta\delta^3, \ldots \right\} \). The parameter \( \delta \) represents standard time consistent impatience. The parameter \( \beta \) represents the self-control problem of the agent where smaller \( \beta \) means the agent has more significant self-control problems. In other words, \( \beta \) is called as the time inconsistent preference for immediate gratification. For \( \beta \equiv 1 \), agent has completely time consistent preferences. The only difference between SHA and NHA is that NHA is not aware of her future preference reversals or self-control problems. SHA is fully aware of her self-control problems so that she predicts how her future selves will behave in the future correctly. Thus, NHA thinks that she will evaluate future payoffs with discount factor \( \delta \) but, in fact, she will evaluate them with \( \beta\delta \).
whereas SHA correctly predicts that she will evaluate future payoffs with $\beta \delta$.

The model is as follows. Time is discrete and the time horizon is infinite. There is one principal and one agent, e.g. a hospital and a doctor to be hired as in Example-1. Along the exogenously given $\hat{T}$ periods, the agent invests to reach some amount of human capital (also exogenously given). The agent faces a 0 or 1 investment decision to make at each time $t$. This implies a fixed cost, $C$, for each unit of investment. In order to finish the investment phase, $k$ units of investment have to be made where $k \leq \hat{T}$. $T$ is the time at which the investment is finished and satisfies $k \leq T \leq \hat{T}$. Provided that the agent finishes the investment phase, the second stage game is played that is bargaining between the agent and the principal, which ends at the period in which it is played in equilibrium. The pie that will be shared in the bargaining game is the amount of human capital that the agent accumulated, $k$ units of investment. Define $w$ as the equilibrium wage determined in the bargaining stage of the game that will be earned at and after $T$ continuously.

All agents will work backwards such that they will predict the wage -EA and SHA predicts correctly but NHA is mistaken in her prediction- that they will earn and depending on this wage earnings, they have to decide on the distribution of investments ($k$ units of investment in $\hat{T}$ periods) and automatically the time they finish the first phase.

By using the above specification of the model, we get the following results:

- In the alternating offer bargaining game where the principal (EA) makes the first offer, under the assumption of both NHA and SHA has same $\beta$ and $\delta$, NHA always gets strictly more payoff than SHA.

If we have a homogeneous cost structure and a no-partial-reward system in the investment stage and assume optimality of finishing the investment stage then,

- Agents always invest consecutively regardless of their preferences.

- The naive agent finishes the investment stage without any delay whereas Sophisticated agent has a periodical investment schedule along the time path-SHA’s belief about her investment behavior is cyclical. She will expect to start the project at every $t_k^*$ periods where $t_k^*$ is the maximum tolerable delay time and it will be defined formally.
• Depending on the parametric values, investment schedule of SHA is found.

• The existence of a specific value for $\beta$ -preference for immediate gratification- that makes SHA finish investment stage without delay is shown. Analogously, the existence of a specific value for $C$ -homogeneous cost of each unit of investment- that makes SHA finish investment stage without delay is shown.

If we have a bonus scheme in the investment stage and the project is not worthwhile for the SHA then:

• The bonus should increase in order for NHA to continue to invest and finish the investment stage. Otherwise, she will procrastinate.

• The agents with higher self-control problems - lower $\beta$'s- should be given higher bonus, with a small caveat, by the principal in order to induce them to complete the same investment project.

The rest of the paper is organized as follows. Section 2 describes the formal model in detail. Section 3 considers the subgame perfect equilibrium of the considered game. Section 4 provides some extensions of the existing model. Section 5 concludes the paper with a brief discussion of the results.

### 3.2 The Formal Model

We suppose that the principal is exponential discounter and the agent is either exponential agent (EA, time consistent) or naive hyperbolic agent (NHA) or sophisticated hyperbolic agent (SHA). The environment is as follows. Time is discrete and the time horizon is infinite. There is one principal and one agent. Along the exogenously given $\hat{T}$ periods, the agent invests to reach to some amount of human capital $V$ (exogenously given also). For simplicity, the agent faces with a 0 or 1 investment decision, $e_t$, at each time $t$, $e_t = 0$ or $e_t = 1$.

Provided that the agent finishes the investment phase, the second stage game is played that is bargaining between the agent and the principal. In order to finish the investment phase, $k$ units of investment have to be made. If we call the marginal value of investment as $f(e_t)$, where $f(0) = 0$, then, since the agent can only make
0 or 1 unit of investment in each period, we can write

\[ V = kf(1) \]

where \( V \) is the value of human capital that is required for the job.

In this framework, investments are costly tasks as reading journals, attending courses, training programs...etc. The cost can be interpreted as the opportunity cost of having spent the time on these activities or the disutility of these activities to the agent. Human capital investment accumulates over time.

We can suppose \( \hat{V} \) is exogenously given because, e.g., the basic requirements or skills (or self-investment) needed to be acquired are announced by the principal. \( V \) is basically the size of the pie to be shared between the principal and the agent after the agent’s \( T \leq \hat{T} \) period investment phase. Since we know the pie size and the types of the players are revealed in the bargaining stage, we can find the subgame perfect equilibrium (SPE) of it. Then, by knowing the reward that will be earned, agents can decide on their investment distribution in the first stage, \( \{e_t\}_{t=0}^T \).

Along the paper, unless others are indicated, the following assumptions will be made:

- Whenever \( V \) is completed, bargaining game starts (and ends at that period, in equilibrium)
- After \( T \) periods, the agent does not have to make any investment.
- No depreciation on accumulated capital.
- Outside option for both players is zero.
- Accumulating capital does not give any utility other than the expected future wage income, which is determined by the bargaining game.

Now we will make some definitions. We call marginal cost of investment as \( C(e_t) \) and it is time independent, strictly convex and satisfies \( C(0) = 0 \). Since we allow only zero or one unit investment choice, for notational convenience, let us define \( C(1) = C \). The value of human capital accumulated up to time \( t \) is \( V_t \).
$V_t$ is a step function and it is weakly increasing between 0 and $T$. After $T$, it is constant. The relationship between $V_t$ and $e_t$ can be written as:

$$\hat{V}_{t+1} - \hat{V}_t = f(e_t) \quad \text{or} \quad \hat{V}_{t+1} = f(e_0) + f(e_1) + \ldots + f(e_t) = \sum_{j=0}^{t} f(e_j). \quad (3.1)$$

Indeed, if the agent finishes the project, then:

$$\hat{V} = \hat{V}_T = \sum_{j=0}^{T-1} f(e_j) = f(e_0) + f(e_1) + \ldots + f(e_{T-1}) \quad (3.2)$$

Define $k$ as the amount of investment required to be completed. The following means the agent finishes the first stage within the time constraint:

$$\sum_{j=0}^{T-1} e_j = k. \quad (3.3)$$

For convenience, we suppose that

$$k \leq \hat{T}$$

which means, if the agent wishes, she can finish the investment phase. In other words, the required investment is doable in $\hat{T}$ periods.

Define $T$ as the time at which the investment is finished

$$\hat{V} = kf(1) = \sum_{j=0}^{T-1} f(e_j). \quad (3.4)$$

$T$ also satisfies the following:

$$k \leq T \leq \hat{T} \quad (3.5)$$

which means she can finish the investment phase at least in $k$ periods without any delay or at most in $T$ period.

If the following is the case, then it means that the agent completes the required investment amount in the maximum amount of time $T$ and continues to
the bargaining game:

\[ \hat{V} = \sum_{j=0}^{\hat{T}-1} f(e_j). \] (3.6)

If the required investment amount is not completed, the agent cannot continue to the second stage and gets zero payoff.

We define \( w(\hat{V}) \) as the equilibrium wage determined in the bargaining stage of the game. Since the value of \( V \) is given and we can find the equilibrium partition of the bargaining game, we know the equilibrium wage \( w(\hat{V}) \).

The problem can be summarized as follows, at time 0, the agent maximizes her discounted utility subject to her time constraint:

\[
\max_{\{e_t\}_{t=0}^{\hat{T}-1}} \left\{ \beta \delta^{T+1} \frac{w(V)}{1-\delta} - \left[ C(e_0) + \beta \sum_{j=0}^{\hat{T}-1} \delta^j C(e_j) \right] \right\} \] (3.7)

subject to

\[ \sum_{j=0}^{\hat{T}-1} f(e_j) = \hat{V} \quad \text{and} \quad k \leq T \leq \hat{T} \]

We, now, briefly mention how the game proceeds in time. First \( T \) periods, agent plays the self-investment game. At period \( T \), the bargaining game is played and it ends in period \( T \) with equilibrium wage for the agent. From \( T \) onwards (including \( T \)), the agent will get the equilibrium wage at each period. Now the problem can be defined as the following. Given \( f(e_t), C(e_t) \) and \( V \), agent will decide on at each time how much to invest on himself to complete the \( V \) in order to maximize his expected payoff. In other words, the problem is to choose \( \{e_t\}_{t=0}^{\hat{T}-1} \) sequence to maximize the expected payoff such that \( \hat{V} = \sum_{j=0}^{\hat{T}-1} f(e_j) \) supposing the agent being exponential discounter, naive hyperbolic discounter or sophisticated hyperbolic discounter. The first stage of this game is, in some sense, similar to the Admati-Perry’s joint project investment framework but the difference is that the investment here is one-sided and made not by different players but by exponential and hyperbolic agent’s selves.
3.3 Characterizing Equilibrium

In this section, we will characterize the equilibrium of the defined game. There are two stages, namely, self-investment stage and bargaining stage in order. We will find the equilibrium by working backwards since this is a perfect information game.

3.3.1 Second Stage Game

In the second chapter, section 4, this stage of the game is examined and the equilibrium is found in more general cases covering the case we think here. We assume that the principal, e.g. hospital, is an exponential type and the agent, e.g. doctor, is either naive or sophisticated hyperbolic type. We refer to chapter 2.4 for the equilibrium of this stage of the game and proceed directly to the first stage of the game.

3.3.2 First Stage Game

We now find the equilibrium path sequence of investment levels for the agents. Let the equilibrium sequence investment levels be \( \{\hat{c}_i\}_{i=0}^{T-1} \) and \( \hat{c}_i \geq 0, \ i = 1, 2, \ldots, T - 1 \). Also, \((t_1; t_2; \ldots; t_k)\) means that agent invests at \( t_1, t_2, \ldots, t_{k-1} \) and finishes at \( t_k \).

We assume that in second stage bargaining game the principal makes the first offer. Since the bargaining game is ahead at least \( k \) periods from now on, NHA will think that she will behave consistently at the bargaining game and she predicts the outcome of it according to the perception that she and the principal have same preferences, \( \delta \). So, from the second stage game, the expected wage of NHA would be

\[
 w_{NHA}(\hat{V}) = \hat{V} \frac{\delta}{1 + \delta}
\]

However, when she actually finishes the first stage (if she does so) and go on bargaining game, she will get \( \beta w_{NHA}(\hat{V}) \). On the other hand, SHA predicts the true wage as it will be in the bargaining game. Expected (and realized) wage of SHA would be

\[
 w_{SHA}(\hat{V}) = \hat{V} \frac{\beta \delta(1 - \delta)}{1 - \beta \delta^2}
\]
Similarly, expected wage of EA would be

\[ w_{NHA}(\hat{V}) = \hat{V} \frac{\delta}{1 + \delta} \]

\[ w_{EA}(\hat{V}) = w_{NHA}(\hat{V}) > w_{SHA}(\hat{V}) \]

We also make an assumption that finishing the first stage game is optimal for all agents, that is, SHA finds it optimal to finish the game (this implies it is optimal for NHA and EA too).

Assumption: Completion of the investment phase of the game is optimal:

\[ (0, 1, 2, \ldots, k) = \left( \frac{\beta \delta^k}{1 - \delta} \right) w_{SHA}(\hat{V}) - \beta \sum_{j=1}^{k-1} \delta^j C - C > 0 \]

This implies optimality for NHA and implies the following too, for EA:

\[ (0, 1, 2, \ldots, k) = \left( \frac{\delta^k}{1 - \delta} \right) w_{EA}(\hat{V}) - \sum_{j=0}^{k-1} \delta^j C > 0 \]

One implication of the above assumption is the following: For NHA, finishing investment by starting immediately and investing consecutively is always better than postponing this one period for any amount of investment left. In other words, the following assumption

\[ \frac{\beta \delta^k}{1 - \delta} w_{SHA}(\hat{V}) - \beta \delta^{k-1} C - \beta \delta^{k-2} C - \ldots - \beta \delta C - C > 0 \]

or \((t_1; t_1 + 1; \ldots; t_1 + k - 1) > 0\)

implies that for NHA

\((t_1; t_1 + 1; \ldots; t_1 + j - 1) > (t_1 + 1; t_1 + 2; \ldots; t_1 + j) \quad \forall j \leq k\)

This can be shown as follows: assume that
\[
\frac{\beta\delta^k}{1 - \delta} w_{SHA}(\hat{V}) - \beta\delta^{k-1} C - \beta\delta^{k-2} C - \ldots - \beta\delta C - C > 0
\]

or \((t_1; t_1 + 1; \ldots; t_1 + k - 1) > 0 \Rightarrow \)
\[
\frac{\beta\delta^k}{1 - \delta} \hat{V} \left(1 + \beta \delta + \ldots + \beta \delta^{k-2} + \beta \delta^{k-1}\right) (1 - \beta \delta^2) > C \Rightarrow (3.8)
\]
\[
\frac{\beta^2 \delta^{k+1} \hat{V}(1 - \delta)}{(1 + \beta \delta - \delta - \beta \delta^k)(1 - \beta \delta^2)} > C
\]

Now, check whether the following is true:

\((t_1; t_1 + 1; \ldots; t_1 + j - 1) > (t_1 + 1; t_1 + 2; \ldots; t_1 + j) \quad \forall j \leq k\)

check first for \(j = k\):

\[
\frac{\beta\delta^k}{1 - \delta} w_{NHA}(\hat{V}) - \beta\delta^{k-1} C - \ldots - \beta\delta C - C > 0
\]

or \(\frac{\beta\delta^{k+1}}{1 - \delta} w_{NHA}(\hat{V}) - \beta\delta^k C - \ldots - \beta\delta C > (3.9)\)

If (3.9) is satisfied for \(j = k\), then it is satisfied for all \(j < k\), since it is decreasing in \(k\). We now will show that the assumption implies (3.9). Last inequality in (3.8) can be written as follows:

\[
\frac{\beta\delta^{k+1} \hat{V}}{(1 - \beta \delta^k)(1 + \delta)} \left(1 + \beta \delta - \delta - \beta \delta^k\right)(1 - \beta \delta^2) > C
\]

(3.10)

In the above equation, \((a)\) is same with expression (3.8). The condition \((b) < 1\) is sufficient for \((a) > C\) to be satisfied, which is the desired result. Fortunately, it is easy to show that \((b) < 1 \forall k > 1\). For \(k = 1\), it is satisfied too since for NHA,

\((t_2; t_2 + 1) \geq (t_2 + 1; t_2 + 2) \Rightarrow (t_2 + 1) > (t_2 + 2)\)

In order to solve this perfect information game, we will use backward induction
by just starting from the period that one unit of investment is left.

**Step 1:** Suppose we have at time $t_1$ ($k - 1 \leq t_1 \leq \hat{T} - 2$), which is the first time that only one unit of investment is needed to finish the first stage - for $t_1 = \hat{T} - 1$, every agent finishes the investment right away by assumption. She can either finish by investing one unit or she postpones investing one period and finishes it next period. Let’s examine what the agent does when she has different preferences:

1a. **Exponential Agent (EA):** If the agent has exponential discounting, then she has two different choices that she should make a decision on. Either she invests now, $t_1$, and finishes it or she postpones investing and finishes it next period. If she invests now, she gets the payoff:

$$-C + \delta w_{EA}(\hat{V}) + \delta^2 w_{EA}(\hat{V}) + ... = \frac{\delta}{1 - \delta} w_{EA}(\hat{V}) - C = \frac{\delta^2}{1 - \delta^2} \hat{V} - C \quad (3.11)$$

If she postpones investment one period and then finishes, she gets:

$$\frac{\delta^3}{1 - \delta^2} \hat{V} - \delta C \quad (3.12)$$

As easily seen,

$\delta(3.11) = (3.12) > 0$ implies $(3.11) > (3.12) > 0$. It is also obvious that postponing more than one period is not optimal either. Thus, EA chooses to finish the investment phase right away when only one unit of investment left.

1b. **Naive Hyperbolic Agent (NHA):** NHA compares two different options like EA such that

$$0 \begin{array}{c} \leq \True \end{array} \frac{\beta \delta^2}{1 - \delta^2} \hat{V} - C \quad (3.13)$$

and
As the assumption implies, finishing right away is better than postponing one period or \((t_1) > (t_1 + 1)\). Thus, NHA finishes investment whenever one unit of investment left.

**1c. Sophisticated Hyperbolic Agent (SHA):** What SHA does is a little different from NHA. We will work backwards. We know that if SHA is at \(t_1 = \hat{T} - 1\), then she will invest for sure. Let’s define a critical value for delaying time like the following:

\[
t^*_1 = \min\{s \in \{1, 2, \ldots\}| \beta \delta w_{SHA}(\hat{V}) - C \geq \frac{\beta \delta^{s+1}}{1 - \delta} w_{SHA}(\hat{V}) - \beta \delta^s C\}
\]

This means that SHA can tolerate \(t^*_1 - 1\) periods of delay. In other words,

\[
\frac{\beta \delta^2}{1 - \delta} w_{SHA}(\hat{V}) - \beta \delta C > \frac{\beta \delta^3}{1 - \delta} w_{SHA}(\hat{V}) - \beta \delta^2 C > ... > \frac{\beta \delta^{t^*_1}}{1 - \delta} w_{SHA}(\hat{V}) - \beta \delta^{t^*_1-1} C > \frac{\beta \delta}{1 - \delta} w_{SHA}(\hat{V}) - C > \frac{\beta \delta^{t^*_1+1}}{1 - \delta} w_{SHA}(\hat{V}) - \beta \delta^{t^*_1} C
\]

or

\[
(t_1 + 1) > (t_1 + 2) > ... > (t_1 + t^*_1 - 1) > (t_1) > (t_1 + t^*_1)
\]

Note that we are at period \(t_1\) and above expression means that delaying \(t^*_1 - 1\) period is acceptable but \(t^*_1\) is not. We can now continue our analysis. Suppose that SHA is at \(t_1 = \hat{T} - 2\). Since she knows that if she postpones to \(\hat{T} - 1\) then she will finish it for sure. Since one period delay is acceptable, then SHA will postpone it one period. Suppose that SHA is at \(t_1 = \hat{T} - 3\). Since she knows that if she postpones to \(\hat{T} - 2\) then she will finish it at \(\hat{T} - 1\). Since two period delay is acceptable, then SHA will postpone it to \(\hat{T} - 2\) and then to \(\hat{T} - 1\) and finish at \(\hat{T} - 1\).
If we make this $t_1^* - 1$ times, SHA again will postpone because $t_1^* - 1$ period delay is acceptable. Now, suppose SHA is at period $t_1 = \hat{T} - 1 - t^*_1 = \hat{T} - t^*_1 - 1$. Then, she knows that if she postpones investing now then she will invest at $\hat{T} - 1$ and it is not optimal to postpone $t^*_1$ periods which means $(\hat{T} - t^*_1 - 1) > (\hat{T} - 1)$. Thus, she invests at $t_1 = \hat{T} - t^*_1 - 1$ and finish the investment phase. Suppose $t_1 < \hat{T} - t^*_1 - 1$. The above argument can be repeated that she will postpone if $\hat{T} - 1 - 2t^*_1 < t_1 < \hat{T} - t^*_1 - 1$. However, if $t_1 = \hat{T} - 1 - 2t^*_1$, then she invests.

Notice that we can go on this iteration and we can say that SHA has a periodically structured investment plan. The behavior characterizations of SHA and other agents are established in the following result.

**RESULT 1:** Suppose Agents are at time $t_1$ and there is only one unit of investment left. Also suppose that completion of the investment phase of the game is optimal. Then, For all parametric values, $\beta, \delta, \hat{T}, \hat{V}$ and for all functional forms $w(\hat{V}), f(,), C(,);$

- $\leftrightarrow$ **EA:** She always finishes the first stage immediately,
- $\leftrightarrow$ **NHA:** She always finishes the first stage immediately,
- $\leftrightarrow$ **SHA:** She invests and finish the first stage at $t_1$ if and only if $t_1 \in \{\hat{T} - 1 - it^*_1\}$ where $i \in \{0, 1, 2, \ldots\}$ such that $0 \leq \hat{T} - 1 - it^*_1$. If $t_1 \notin \{\hat{T} - 1 - it^*_1\}$, then she will postpone investing up to the closest time such that $t_1 \in \{\hat{T} - 1 - it^*_1\}$. If $\hat{T} - 1 - t^*_1 < 0$, then SHA follows the $(\hat{T} - 1)$ strategy.

**Proof:**

The argument is in the above substeps, $1a, 1b, 1c$. The explanation for the last part is that If $\hat{T} - 1 - t^*_1 < 0$ or $\hat{T} - 1 < t^*_1$ then the available time for investment is, $\hat{T} - 1$, less than or equal to the maximum tolerance time, $t^*_1 - 1$, which means she can tolerate to postpone the investment up to the period $\hat{T} - 1$. Thus, she follows $(\hat{T} - 1)$ strategy.

**Step 2:** Suppose we are at time $t_2$, which is the first time that only two units of investment are needed to finish the first stage. She can either invest one unit and leaves one unit to finishing or she postpones investing. Let’s again examine what the agent does when she has different preferences:
2a. Exponential Agent (EA): Suppose we are at $t_2$ such that $k - 2 \leq t_2 \leq \hat{T} - 2$. What the EA can do is that she can either invest now and finish it next period (from step 1) or she can postpone investing $\tilde{t}$ period at $t_2$ to $t_2 + \tilde{t}$ and then finish it at $t_2 + \tilde{t} + 1$ where $0 < \tilde{t} \leq \hat{T} - t_2 - 1$. Let’s compare the payoffs:

$$\frac{\delta^2}{1 - \delta} w_{EA} (\hat{V}) - \delta C - C > \frac{\delta^{\tilde{t} + 1}}{1 - \delta} w_{EA} (\hat{V}) - \delta^{\tilde{t} + 1} C - \delta^{\tilde{t}} C$$

(3.14)

So, for any time, whenever two units of investment needed to finish, the EA finishes it in two periods by investing consecutively.

2b. Naive Hyperbolic Agent (NHA): Suppose we are at $t_2$ such that $k - 2 \leq t_2 \leq \hat{T} - 2$. NHA compares different options like the following:

$$
\begin{align*}
(t_2; t_2 + 1) &\quad (t_2 + 1; t_2 + 2) \quad \dot{\ldots} \quad (\hat{T} - 2; \hat{T} - 1) \\
(t_2; t_2 + 2) &\quad (t_2 + 1; t_2 + 3) \quad \dot{\ldots} \quad (\hat{T} - 3; \hat{T} - 1) \\
(t_2; t_2 + 3) &\quad \dot{\ldots} \\
(t_2 + 1; \hat{T} - 1) &\quad (t_2 + 1; \hat{T} - 1) \\
(t_2; \hat{T} - 1) &\quad (t_2; \hat{T} - 1)
\end{align*}
$$

(3.15)

The number of options that she has to consider is a lot ($ (T - t_2 - 1)(T - t_2)/2 $) but indeed, when we compare the payoffs of them, we recognize the fact that choosing the investment periods close to each other is better than having more periods between investments, e.g., $(t_2 + 2; t_2 + 3) > (t_2 + 1; t_2 + 3) > (t_2; t_2 + 3)$ and that making the investments as close as possible is better than postponing more and more, e.g., $(t_2 + 1; t_2 + 2) = \frac{1}{\delta}(t_2 + 2; t_2 + 3) = \frac{1}{\delta^2}(t_2 + 3; t_2 + 4)$. So, only options that should be compared are $(t_2; t_2 + 1)$ and $(t_2 + 1; t_2 + 2)$. We know that

$$(t_2; t_2 + 1) \geq (t_2 + 1; t_2 + 2)$$
She invests at $t_2$. When tomorrow, $t_2 + 1$, comes, she will solve the problem in Step 1. But

$$(t_2; t_2 + 1) \geq (t_2 + 1; t_2 + 2) \Rightarrow (t_2 + 1) > (t_2 + 2)$$

Thus, she will follow $(t_2; t_2 + 1)$ strategy.

2c. Sophisticated Hyperbolic Agent (SHA): There are two cases:

If $(t_2; t_2 + 1) \geq (t_2 + 1; t_2 + 2)$ or

$$\frac{\beta \delta^2}{1 - \delta} w_{SHA}(\hat{V}) - \beta \delta C - C \geq \frac{\beta \delta^3}{1 - \delta} w_{SHA}(\hat{V}) - \beta \delta^2 C - \beta \delta C$$

invest at $t_2$ and $t_2 + 1$

Then, since $(t_2; t_2 + 1) \geq (t_2 + 1; t_2 + 2) \Rightarrow (t_2 + 1) > (t_2 + 2)$, she will finish it by investing consecutively, $(t_2; t_2 + 1)$.

If $(t_2; t_2 + 1) < (t_2 + 1; t_2 + 2)$, she knows that if she postpones one period, she will keep postponing up to $\hat{T} - 2$. To solve SHA’s problem, we will work backwards again. We know that if SHA is at $t_2 = \hat{T} - 2$, then She will invest for sure. Suppose $t_2 = \hat{T} - 3$. Then SHA has two options $(\hat{T} - 3; \hat{T} - 1)$ and $(\hat{T} - 2; \hat{T} - 1)$, because she knows that if she invest at $\hat{T} - 3$, she will invest the last unit at $\hat{T} - 1$ from 1c. Also, if she postpones investment to next period, she will finish it for sure by assumption. Thus, SHA postpones investment. Suppose $t_2 = \hat{T} - 4$. Then, by the same argument, she will compare the following options: $(\hat{T} - 4; \hat{T} - 1)$, $(\hat{T} - 3; \hat{T} - 1)$ and $(\hat{T} - 2; \hat{T} - 1)$. The last option is optimal for him, so SHA postpones investment two periods. This iteration goes on up to the period where $t_2 = \hat{T} - 2 - t_1^*$. At $t_2 = \hat{T} - 2 - t_1^*$, if SHA invests then she will invest the last unit at $\hat{T} - 2 - t_1^*$ by Result 1. Then SHA will compare the following two options $(\hat{T} - 2 - t_1^*; \hat{T} - 1 - t_1^*)$ and $(\hat{T} - 2; \hat{T} - 1)$.

If $(\hat{T} - 2 - t_1^*; \hat{T} - 1 - t_1^*) \geq (\hat{T} - 2; \hat{T} - 1)$, or

$$\frac{\beta \delta^2}{1 - \delta} w_{SHA}(\hat{V}) - \beta \delta C - C \geq \frac{\beta \delta^{t_1^*+2}}{1 - \delta} w_{SHA}(\hat{V}) - \beta \delta^{t_1^*+1} C - \beta \delta^{t_1^*} C$$

(3.16)

Then, SHA will finish it immediately.
We can repeat exactly the same argument above. Now, SHA knows that she will follow $(\hat{T} - 2 - t_1^*; \hat{T} - 1 - t_1^*)$ strategy. For $t_2 \in \{\hat{T} - 3 - t_1^*, \hat{T} - 4 - t_1^*, ..., \hat{T} - 1 - 2t_1^*\}$, she keeps postponing up to the period $t_2 = \hat{T} - 2 - t_1^*$. At $t_2 = \hat{T} - 2 - 2t_1^*$, she will compare $(\hat{T} - 2 - 2t_1^*; \hat{T} - 1 - 2t_1^*)$ and $(\hat{T} - 2 - t_1^*; \hat{T} - 1 - t_1^*)$. Notice that this comparison is the same with the one in expression $(2c - 1)$. So, she decides to finish it right away, $(\hat{T} - 2 - 2t_1^*; \hat{T} - 1 - 2t_1^*)$. Like in Step1, we can go on this iteration and we can say that SHA has a periodically structured investment plan.

If $(\hat{T} - 2 - t_1^*; \hat{T} - 1 - t_1^*) < (\hat{T} - 2; \hat{T} - 1)$, Then, she will postpone. Remember we are at time $t_2 \leq \hat{T} - 2 - t_1^*$. We define a critical value for second level delaying time as follows:

$$t_2^* = \min\{s \in \{j(t_1^*)\} \mid \frac{\beta \delta^2 w_{SHA}(\hat{V})}{1 - \delta} - \beta \delta C - C \geq \frac{\beta \delta^{s+2}}{1 - \delta} w_{SHA}(\hat{V}) - \beta \delta^{s+1} C - \beta \delta^s C\}$$

$j \in \{1, 2, ...\}$ such that $0 \leq \hat{T} - 1 - (j - 1)t_1^*$.

This definition implies that $t_2^* - 1$ is a tolerable amount of time if there are two units of investment but $t_2^*$ is not tolerable.

Then SHA will follow $(\hat{T} - 2 - t_1^*; \hat{T} - 1 - t_1^*)$. Since She knows this, she will have again a periodically structured strategy. She will postpone the investment to $\hat{T} - 2 - t_2^*$ if $\hat{T} - 2 - 2t_1^* < t_2 < \hat{T} - 2 - t_2^*$. Moreover, if $t_2 = \hat{T} - 2 - 2t_1^*$, she will invest immediately. Similarly, we can continue iteration like this. The behavior characterizations of SHA and other agents are established in the following result.

We can repeat exactly the same argument above. Now, SHA knows that she will follow $(\hat{T} - 2 - t_1^*; \hat{T} - 1 - t_1^*)$ strategy. For $t_2 \in \{\hat{T} - 3 - t_1^*, \hat{T} - 4 - t_1^*, ..., \hat{T} - 1 - 2t_1^*\}$, she keeps postponing up to the period $t_2 = \hat{T} - 2 - t_1^*$. At $t_2 = \hat{T} - 2 - 2t_1^*$, she will compare $(\hat{T} - 2 - 2t_1^*; \hat{T} - 1 - 2t_1^*)$ and $(\hat{T} - 2 - t_1^*; \hat{T} - 1 - t_1^*)$. Notice that this comparison is the same with the one in expression $(2c - 1)$. So, she decides to finish it right away, $(\hat{T} - 2 - 2t_1^*; \hat{T} - 1 - 2t_1^*)$. Like in Step1, we can go on this iteration and we can say that SHA has a periodically structured investment plan.

If $(\hat{T} - 2 - t_1^*; \hat{T} - 1 - t_1^*) < (\hat{T} - 2; \hat{T} - 1)$, Then, she will postpone. Remember we are at time $t_2 \leq \hat{T} - 2 - t_1^*$. We define a critical value for second level delaying time as follows:
\[ t_2^* = \min\{ s \in \{ j(t_1^*) \} | \frac{\beta \delta^2 w_{SHA}(\hat{\mathcal{V}})}{1 - \delta} - \beta \delta C - C \geq \frac{\beta \delta^2+2}{1 - \delta} w_{SHA}(\hat{\mathcal{V}}) - \beta \delta^2+1 C - \beta \delta C \} \]

such that \( j \in \{1, 2, \ldots\} \) such that \( 0 \leq \hat{T} - 1 - (j - 1)t_1^* \).

(notice that the above case, \( 2c - 1 \), is just \( s = t_2^* = t_1^* \)). This definition implies that \( t_2^* - 1 \) is tolerable amount of time if there are two units of investment but \( t_2^* \) is not tolerable.

Then SHA will follow \((\hat{T} - 2 - t_2^*; \hat{T} - 1 - t_2^*)\). Since She knows this, she will have again a periodically structured strategy. She will postpone the investment to \( \hat{T} - 2 - t_2^* \) if \( \hat{T} - 2 - 2t_2^* < t_2 < \hat{T} - 2 - t_2^* \). Moreover, if \( t_2 = \hat{T} - 2 - 2t_2^* \), she will invest immediately. Similarly, we can continue the iteration like this. The behavior characterizations of SHA and other agents are established in the following result.

**RESULT 2:** Suppose Agents are at time \( t_2 \) and there are two units of investment left. Also suppose that completion of the investment phase of the game is optimal. Then, For all parametric values, \( \beta, \delta, \hat{T}, \hat{\mathcal{V}} \) and for all functional forms \( w(\hat{\mathcal{V}}), f(.), C(.) \);

\( \hookrightarrow \) EA: She always finishes the first stage immediately,

\( \hookrightarrow \) NHA: She always finishes the first stage immediately,

\( \hookrightarrow \) SHA: She finishes immediately, e.g. \((t_2; t_2 + 1)\), if and only if \( t_2 \in \{ \hat{T} - 2 - it_2^* \} \) where \( i \in \{0, 1, 2, \ldots\} \) such that \( 0 \leq \hat{T} - 2 - it_2^* \). If \( t_2 \notin \{ \hat{T} - 2 - it_2^* \} \), then she will postpone investing up to the closest time such that \( t_2 \in \{ \hat{T} - 2 - it_2^* \} \). If \( \hat{T} - 2 - t_2^* < 0 \), then SHA follows the \((\hat{T} - 2; \hat{T} - 1)\) strategy.

**Proof:**

The argument is in the above substeps, \( 2a, 2b, 2c \). The explanation for the last part is that If \( \hat{T} - 2 - t_2^* < 0 \) or \( \hat{T} - 2 < t_2^* \) then the available time for investment is, \( \hat{T} - 2 \), less than or equal to the maximum tolerance time, \( t_2^* - 1 \), which means she can tolerate to postpone the investment up to the period \( \hat{T} - 2 \). Thus, she follows \((\hat{T} - 2; \hat{T} - 1)\) strategy.
We can continue these steps and at each time $t_i$ that $i$ units of investment left, find a critical value $t_i^*$ for $t_i$. We are going to use the method of induction in order to show that this is true.

**Theorem 1:** Suppose Agents are at time $t_k \geq 0$ and there are $k$ units of investment needed to finish the first stage. Also suppose that completion of the investment phase of the game is optimal. Define $t_k^*$ as follows:

$$t_k^* = \min \{ s \in \{ j(t_{k-1}^*) \} \beta^j \delta^{k+i} w_{SHA}(\hat{V}) \frac{1}{1-\delta} - \beta^{k-1} \delta^i C - C \geq \frac{\beta \delta^{s+k} w_{SHA}(\hat{V})}{1-\delta} - \beta_{j=0}^{k-1} \delta^{s+i} C \}$$

$j \in \{1, 2, \ldots \}$ such that $0 \leq \hat{T} - k - (j - 1)(t_{k-1}^*)$.

Then, For all $\beta, \delta, \hat{T}, \hat{V}$ and for all functional forms $w(\hat{V}), f(.), C(.)$;

$\rightarrow$EA: She, always, finishes the first stage without any delay,

$\rightarrow$NHA: She, always, finishes the first stage without any delay,

$\rightarrow$SHA: She finishes immediately, e.g. $(t_k; t_k + 1; \ldots; t_k + k - 1)$, if and only if $t_k \in \{ \hat{T} - k - it_k^* \}$ where $i \in \{0, 1, 2, \ldots \}$ such that $\hat{T} - k - it_k^* \geq 0$.

If $\{ \hat{T} - k - it_k^* \} \notin t_k$, then she will postpone investing up to the closest time such that $t_k \in \{ \hat{T} - k - it_k^* \}$. If $\hat{T} - k - t_k^* < 0$, then SHA follows the $(\hat{T} - k; \hat{T} - k + 1; \ldots; \hat{T} - 2; \hat{T} - 1)$ strategy.

**Proof:**

First, we will mention some important points about the investment schedule of the agents. As is explained above, no agent want to leave any time gaps between investment periods. The first reason is the homogeneous cost structure in the investment game and that there is no uncertainty about the cost that the agent should pay at each period if she invests. This makes agents certain about their contingent plans for the future. The second reason is that the reward system is constructed in such a way that there is no partial or gained utility unless the investment is completely finished. Leaving time gap between investment periods
always makes them worse off. Given the completion time, all agents want to make the costly investments as close as possible to the completion period in order to minimize the cost since there is discounting.

For EA, the argument is same as above steps. For her, postponing is never optimal because of her exponential discounting type. Thus, she finishes the first stage without any delay.

For NHA, she will again compare only the options of finishing immediately and postponing one period as follows:

\[(t_k; t_k + 1; \ldots; t_k + k - 1) \& (t_k + 1; t_k + 2; \ldots; t_k + k)\]

Since finishing immediately is always optimal for her:

\[(t_k; t_k + 1; \ldots; t_k + k - 1) > (t_k + 1; t_k + 2; \ldots; t_k + k)\]

Thus, she finishes the first stage without any delay.

For SHA, we will use the induction method. The first three steps were explained above. Assume that we have a value for \(t_{k-1}^*\) and show that we have a \(t_k^*\) such that above condition for SHA is satisfied.

The point that \(t_i^*\) is chosen from the multiples of \(t_{i-1}^*\) is important. The reason for this is that, by definition of \(t_{i-1}^*\), SHA knows for sure that she will only invest at periods that are in the following time schedule \(\hat{T} - (i - 1) - j t_{i-1}^*\), if \(i - 1\) units of investment left. Other than those periods, she will postpone investing up to the closest time that is in that time schedule because by investing at \(\hat{T} - i - j t_{i-1}^*\), she minimizes the cost that she will incur. So, she invests the \(i^{th}\) left unit according to the time schedule \(\hat{T} - i - j t_{i-1}^*\). Then, she compares the payoffs of investing for different \(j\) values and she does this according to the maximum tolerable time of postponing that is basically the definition of \(t_k^*\).

As it is explained above SHA invests the \(k^{th}\) unit at \((\hat{T} - k - it_{k-1}^*)\), which is one period before than the periods that she will surely invest the \((k - 1)^{th}\) unit (and the rest too), which are \(\hat{T} - (k - 1) - it_{k-1}^*\) (above figure). In other periods, investing the \(k^{th}\) unit is not optimal because next period she will not invest the rest and she will wait the closest period that is in the time schedule \(\hat{T} - (k - 1) - it_{k-1}^*\). So, she invests at periods \((\hat{T} - k - it_{k-1}^*)\). Then, what she will do is to calculate the
maximum tolerable time for postponing but the definition of \( t^*_k \) gives this. \( t^*_k - 1 \) is the maximum tolerable time for SHA to make the last \( k^{th} \) unit of the investment. Thus, she invests at period \( \hat{T} - k - t^*_k \). With the same logic used in earlier steps, since she is sophisticated, she knows this and at earlier periods than \( \hat{T} - k - t^*_k \), she will take this into account and again will have a periodic investment scheme like \( \hat{T} - k - it^*_k \). Thus, she will invest consecutively whenever she is at the period that is in the time schedule \( \hat{T} - k - it^*_k \).

The explanation for the last part is that If \( \hat{T} - k - t^*_k < 0 \) or \( \hat{T} - k < t^*_k \), then the available time for investment is, \( \hat{T} - k \), less than or equal to the maximum tolerance time, \( t^*_k - 1 \), which means she can tolerate to postpone the investment up to the period \( \hat{T} - k \). Thus, she follows the \( (\hat{T} - k; \hat{T} - k + 1; \ldots; \hat{T} - 2; \hat{T} - 1) \) strategy.

\[ \square \]

**Corollary 1:** The following is always satisfied:

\[ t^*_k \geq t^*_{k-1} \geq \ldots \geq t^*_2 \geq t^*_1 \geq 1 \]

Moreover, if maximum tolerable time is zero for the first investment or \( t^*_k = 1 \), then \( t^*_k = t^*_{k-1} = \ldots = t^*_2 = t^*_1 = 1 \). Thus, SHA finishes the first stage without any delay.

**Proof:**

The reason for the first condition is the following: \( t^*_1 \geq 1 \) by definition. Moreover, for all \( i = k, k - 1, k - 2, \ldots, 2 \), the definition of the \( t^*_i \) entails that any \( t^*_i \) is chosen from the multiples of \( t^*_i - 1 \) or \( t^*_i \in j(t^*_i - 1) \) where \( j = 1, 2, 3 \ldots \). Thus the condition should be satisfied. If \( t^*_k = 1 \), then from the first condition, all other \( t^*_i \)s should be one, too. This can be shown in a different way like the following:

\[ t^*_k = 1 \text{ implies that} \]

\[ \frac{\beta \delta^k w_{SHA}(\hat{V})}{1 - \delta} - \beta^i_{j=1} \delta^j C - C \geq \frac{\beta \delta^{i+1} w_{SHA}(\hat{V})}{1 - \delta} - \beta^i_{j=0} \delta^{i+1} C \implies \]

\[ \frac{\beta \delta^k w_{SHA}(\hat{V})}{1 - \beta \delta^k} \geq C \]
The left side of the above inequality is decreasing in \(k\) implies that the inequality is valid for values smaller than \(k\). Thus, if \(t_k^* = 1\), then all other critical values should be one too.

\[\square\]

**Theorem 2:** For given values of \(\delta, \hat{T}, \hat{V}, k\) and \(C\), \(\exists \beta^*\) such that for all \(\beta \geq \beta^*\), SHA finishes the first stage without any delay.

**Proof:** Assume \(\delta, \hat{T}, \hat{V}, k\) and \(C\) values are given. In order SHA to finish the first stage without delay, the following condition should be satisfied:

\[
(t_k; t_k + 1; \ldots; t_k + k - 1) > (t_k + 1; t_k + 2; \ldots; t_k + k)
\]

or

\[
\frac{\beta \delta^k w_{SHA}(\hat{V})}{1 - \beta \delta^k} \geq C
\]

Now, plug the expression for \(w_{SHA}(\hat{V})\) into the above inequality and get:

\[
\frac{\beta \delta^k \hat{V} \delta (1 - \delta)}{(1 - \beta \delta^2)} \geq C \implies
\]

\[
\frac{\beta^2}{(1 - \beta \delta^2 - \beta \delta^k - \beta^2 \delta^{k+2})} \delta^{k+1}(1 - \delta) \hat{V} \geq C
\]

It is not difficult to show that \(dh(\beta)/d\beta > 0\). In the above equation, if the inequality is actually satisfied with equality then that specific \(\beta\) value is \(\beta^*\). The reason is that when we increase \(\beta\) a little bit, then the left hand-side (LHS) will be larger than RHS, which means SHA still wants to finish first stage without any delay or \(t_k^* = 1\). If we decrease \(\beta\), then the right hand-side (RHS) will be larger than LHS, which means SHA finds it optimal to postpone investment or \(t_k^* > 1\).

Moreover since the following is satisfied:

\[
\frac{\beta \delta^k w_{SHA}(\hat{V})}{1 - \beta \delta^k} \geq C \implies \frac{\beta \delta^j w_{SHA}(\hat{V})}{1 - \beta \delta^j} \geq C, \forall j \leq k
\]

The planned investment schedule is implemented by SHA.

So, there exist a \(\beta^*\) such that
\[
\frac{(\beta^*)^2}{(1 - (\beta^*)\delta^2 - (\beta^*)\delta^k - (\beta^*)^2\delta^{k+2})}\delta^{k+1}(1 - \delta)\hat{V} = C
\]

and for all \( \beta \geq \beta^* \), LHS will be higher than RHS and vice versa. This completes the proof.

We can write the Theorem 2 for the value of cost as well since \( C \) is also a specific characteristic of the agent. It would be like the following: For given values of \( \beta, \delta, \hat{T}, \hat{V} \) and \( k \), \( \exists \ C^* \) such that for all \( C \leq C^* \), SHA finishes the first stage without any delay. The rationale of the Theorem 2 is to point out the role of immediate gratification preference of the agent in the decision making process. It says there exists an immediate gratification preference level, \( \beta \), such that at and above that level, she just starts investing immediately.

In Theorem 2, we assume \( \delta, \hat{T}, \hat{V}, k \) and \( C \) are all given. Actually, values for \( \hat{T}, \hat{V} \) and \( k \) are already exogenous in the problem. Agents take these as given by definition of the problem. For \( \delta \), it is a common discount factor of both agent and the principal and taking it as given is not a very strong assumption. For the cost value, \( C \), it may differ from agent to agent but in this formulation we assume that \( C \) is constant. Later we will assume a different cost scheme that allows more flexibility in the context.

### 3.4 Equilibrium With The Bonus Motive

In this section, we will add a bonus motive to the constructed framework. In order to see the rationale behind this, we remember the examples again. The insurance company worker can get a bonus in her wage by incurring the cost of taking and passing the exam. The doctor can be rewarded by the hospital prior to the actual wage agreement for each of her costly investments. The student can take the courses in each semester and the firm can reward her by giving her the opportunity to be an intern. These can be the examples for having a bonus or this kind of rewarding scheme in this framework. The rationale behind doing this in terms of the company or the hospital is to make the agent not to procrastinate or to finish the investment earlier than the deadline. Because for the principal, waiting is costly in the sense that it cannot earn the profit that it would get in the periods
where agents can finish the investment but instead they procrastinate. Another approach may be that these bonuses may make a job desirable that is actually not worth to finish without bonuses. Thus, as an incentive mechanism, the principal can take advantage of this bonus motive to make the agents finish the investment earlier or to make the investment worth finishing. The bonus structure that will be presented here is different from the one that is mentioned in O’Donoghue and Rabin [2003] as an extension. O’Donoghue and Rabin assumes a fixed total reward scheme causing more severe procrastination but here we will assume that this new reward scheme is, in fact, a ”bonus” in the sense that it is paid extra to the agent by principal without any reduction in the agent’s expected future wage income.

One of the simplest modeling ways of the bonus motive for the firm is to offer a fixed amount of benefit to the agent after each unit of investment is made. Call the bonus amount as $x$. It is earned with a one period lag, e.g., if the agent invests at time $t$ then, she gets $x$ at time $t + 1$. We can also interpret the one-time bonus as the present value of a continuous benefit initiated by the completion of each one unit of investment.

Here, we do not have to assume the optimality of finishing the investment phase of the game because bonus scheme may make an unworthy investment project worthy. In the previous section, the optimality of finishing by SHA implies optimality for others. The interesting case was optimal finishing because if it is not, then they will not even start to project and the game ends. Here again the interesting case is the following: for NHA, it is optimal to finish the project without any delay without the bonus scheme and for SHA, it is not optimal to finish the project without the bonus scheme but optimal to finish without any delay with the bonus scheme. Thus, now we can get an interesting case where SHA has a similar investment structure but NHA may have a procrastinative behavior. Actually, since there is this strategic interaction between the principal and the agent it is not the case that NHA procrastinate inefficiently. In other words, the principal offers a bonus scheme (if it is optimal for herself) that just makes NHA invest consecutively otherwise, if there is no optimal bonus scheme for the principal, she will not introduce it.

So, we assume the followings: For NHA and EA, finishing investment phase in $k$ periods by investing consecutively is optimal;
\[
\frac{\beta \delta^k w_{\text{NHA}}(\hat{V})}{1 - \delta} - \beta \Sigma_{j=1}^{k-1} \delta^j C - C \geq 0
\]

or

\[
\frac{\beta \delta^k w_{\text{NHA}}(\hat{V})}{1 - \beta \delta^k + \beta \delta - \delta} \geq C
\]

(3.17)

However, for SHA it is not;

\[
\frac{\beta \delta^k w_{\text{SHA}}(\hat{V})}{1 - \delta} - \beta \Sigma_{j=1}^{k-1} \delta^j C - C < 0
\]

or

\[
\frac{\beta \delta^k w_{\text{SHA}}(\hat{V})}{1 - \beta \delta^k + \beta \delta - \delta} < 0
\]

(3.18)

We now add the bonus scheme to the model: Equation 3.17 implies the following;

\[
\beta \delta^k [\frac{w_{\text{NHA}}(\hat{V})}{1 - \delta} + x] - \beta \Sigma_{j=1}^{k-1} [\delta^j C - x] - C > 0
\]

(3.19)

We also assume that bonus scheme makes finishing worthy for SHA;

\[
\beta \delta^k [\frac{w_{\text{SHA}}(\hat{V})}{1 - \delta} + x] - \beta \Sigma_{j=1}^{k-1} [\delta^j C - x] - C \geq 0
\]

(3.20)

So 3.17, 3.18 and 3.20 are assumed and 3.19 is indicated by 3.17.

The principal is going to give a bonus to the agents to make them finish the investment phase without delay. So, the Naive agent, e.g., will calculate whether postponing one period is optimal for her or not. It is the following:

\[
\beta \delta^k [\frac{w_{\text{NHA}}(\hat{V})}{1 - \delta} + x] - \beta \Sigma_{j=1}^{k-1} [\delta^j C - x] - C \quad \text{and}
\]

\[
\beta \delta^{k+1} [\frac{w_{\text{NHA}}(\hat{V})}{1 - \delta} + x] - \beta \Sigma_{j=1}^{k-1} [\delta^j C - x] - C
\]

(3.21)
The comparison of these two payoffs is the same with the following comparison:

$$\frac{\beta \delta^k w_{NHA}(\hat{V}) + \beta \delta x(1 - \delta^k)}{1 - \beta \delta^k} \text{ and } C$$

Now we mention a sufficient condition for NHA to finish the investment project right away. By using 3.17:

$$\frac{\beta \delta^k w_{NHA}(\hat{V}) + \beta \delta x(1 - \delta^k)}{1 - \beta \delta^k} \geq \frac{\beta \delta^k w_{NHA}(\hat{V})}{1 - \beta \delta^k + \beta \delta - \delta} \geq C \quad (3.22)$$

If the first inequality is satisfied then, This implies that NHA will start investment at time 0 and invest consecutively and finish it in $k - 1$ periods, in other words, $(0; 1; 2; \ldots; k - 1)$. Now, by using inequality in 3.22, we can find a condition on the bonus, $x$, as follows;

$$x \geq \frac{(1 - \beta)\delta^k w_{NHA}(\hat{V})}{(1 - \delta^k)(1 - \beta \delta^k + \beta \delta - \delta)} \quad (3.23)$$

Notice that, without bonus scheme, NHA will postpone if the following is true:

$$\frac{\beta \delta^k w_{NHA}(\hat{V})}{(1 - \beta \delta^k + \beta \delta - \delta)} \geq C \geq \frac{\beta \delta^k w_{NHA}(\hat{V})}{1 - \beta \delta^k}$$

So, bonus scheme makes NHA follow $(0; 1; 2; \ldots; k - 1)$ strategy if and only if 3.23 is satisfied for all $k$. NHA thinks that if bonus satisfies 3.23 then, she will follow $(0; 1; 2; \ldots; k - 1)$ strategy. However, when the next period comes, NHA again makes the same calculation. She will have the same expression with 3.23 except that instead of $k$ now she has $k - 1$ since she invested the $k^{th}$ unit. Since the right-hand-side in expression 3.23 is decreasing in $k$, the new bonus amount that makes NHA decide to invest is higher now. This may make NHA decide not to invest (it is a weak result because 3.23 is sufficient but not necessary condition.) Thus, as long as 3.23 is satisfied for all $k \geq 1$, then NHA follow $(0; 1; 2; \ldots; k - 1)$ strategy. Otherwise, NHA may start to invest but after some point she may decide not to continue because of the insufficient bonus amount (but since this kind of behavior is obviously not optimal for the principal, she will not allow this to happen by arranging the bonus scheme).
We now think about the principal’s problem. She is going to compare the payoff from the agent and the bonus that will be given to her. She will give the bonus if and only if

\[
\sum_{j=0}^{k-1} \delta_j x_j \leq \hat{V}(1 - \beta \delta) \left( \frac{\delta^k}{1 - \delta} \right) \]  

where (1) is the payoff of principal from hiring the agent per period, (1) * (2) gives the total discounted payoff of principal and (3) is the discounted value of the bonus that she gives to the agent \((x_j's\) will be specified later and it will be shown that they are actually different not the same for all periods for NHA case). If 3.24 is satisfied then, the principal decides to give the bonus.

The problem of NHA is to compare just the immediate cost and the short run benefit of investing that is she only compare the trade of between investing today or tomorrow by assuming once started she will continue investing consecutively. This is a problem because when future investment periods arrive, her immediate gratification may overwhelm her optimistic beliefs of consecutive investment at previous periods and she may give up investing after some point where the bonus is no longer enough to make her continue investing. From the perspective of NHA, whenever condition 3.23 is satisfied, she continues investing and vice versa. Thus, we may well have a situation where NHA starts to invest by having an optimistic beliefs about her future-selves but after sometime she is defeated to her self-control problem and stop investing and this deprives her from getting the wage after investment phase.

However, we have interactions here between agents and bonus scheme will be implemented by the principal if and only if she believes that NHA will finish the investment stage as planned (assume a complete information framework for now, e.g., principal knows the type and the self-control problem of the agent). So, in this kind of environment, either Principal implements the bonus scheme and NHA finishes "efficiently" or she does not implement it and NHA may either not even start or finish it depending on the payoff-cost comparison. In fact, there will not be a case where NHA starts but not finish because of again the homogeneous cost structure and not having immediate rewards (in case of bonus scheme is not
implemented).

The question is which bonus scheme, \( x'_j \)'s, will the principal choose -if there exists a bonus scheme that is optimal for both NHA and the principal? What the principal does is the following: at \( t = 0 \), she is going to promise the agent to give her at least the bonus amount that is:

\[
x_0 = \frac{\hat{V} \beta \delta (1 - \beta) \delta^k}{(1 - \delta^k)(1 + \delta)(1 - \beta \delta^k + \beta \delta - \delta)} \quad (3.25)
\]

The agent makes the trade-off between investing today or postponing one period. This bonus scheme guarantees her a higher payoff if she invests today than postponing-by assuming that she will get at least this much bonus in the future for each of her investment unit. In the next period, the bonus amount will be higher because the expression in 3.25 is decreasing in \( k \). So when \( k - j \) units left, the principal will give the agent:

\[
x_j = \frac{\hat{V} \beta \delta (1 - \beta) \delta^{k-j}}{(1 - \delta^{k-j})(1 + \delta)(1 - \beta \delta^{k-j} + \beta \delta - \delta)} \quad (3.26)
\]

and the agent will accept this bonus and invest, again by assuming that she will get at least this much bonus in the future for each of her unit of investment. Thus the optimal bonus scheme is like in expression 3.26, \( \forall 0 \leq j \leq k - 1 \). This bonus scheme is a sufficient bonus scheme that the principal is willing to give to the agent and the agent is also willing to take it and invest consecutively as she planned at \( t = 0 \).

Now, the existence of the bonus scheme boils down to the following condition:

\[
\hat{V}(1 - \beta \delta)\left(\frac{1}{1 + \delta}\right)(\frac{\delta^k}{1 - \delta}) \geq \sum_{j=0}^{k-1} \delta^j x_j; \quad \forall k
\]

or

\[
\hat{V}(1 - \beta \delta)\left(\frac{1}{1 + \delta}\right)(\frac{\delta^k}{1 - \delta}) \geq \sum_{j=0}^{k-1} \delta^j \frac{\hat{V} \beta \delta (1 - \beta) \delta^{k-j}}{(1 - \delta^{k-j})(1 + \delta)(1 - \beta \delta^{k-j} + \beta \delta - \delta)} \quad (3.27)
\]

or
\[
\left( \frac{1 + \delta - \beta \delta}{\beta \delta (1 - \beta)} \right) \left( \frac{1}{1 - \delta} \right) \geq \sum_{j=0}^{k-1} \frac{1}{(1 - \delta^{k-j})(1 - \beta \delta^{k-j} + \beta \delta - \delta)}
\]

**Proposition 1:** Assume, for given values of \( \delta, \beta \) and \( k \), 3.27 is satisfied (existence of a bonus scheme). Then, in the investment game with bonus scheme, the bonus amount for NHA is a monotonically decreasing function of \( k \). In other words, the bonus should increase in order for the NHA to continue to invest and finish the investment stage.

**Proof:** The argument is above.

The proposition generates a very similar result with the O’Donoghue and Rabin paper [1999b]. In that paper, they find that the optimal incentives for procrastinators typically involve an increasing punishment for delay as time passes. We actually find the complement of this result saying that whenever the agent continues to invest, she should get a higher bonus. These are very similar results because increasing punishment when the agent did not invest is almost the same thing with increasing reward or bonus when they invest. So, we generate the same result in our framework that is more general in the sense that we have more than two investment periods but it is more restricted in the sense that we have a homogeneous cost structure along the investment path. There is this trade-off between having longer projects and having a simpler cost structure. Including both longer projects and more complex cost structure remains to be done.

**Proposition 2:** For the above game, the bonus scheme is an increasing function of self-control problem of the agent, \( \beta \) (with a small caveat). In other words, the agents with higher self-control problems - lower \( \beta \)'s- should be given higher bonus by the principal in order to induce the agents to complete the same investment project.

**Proof:** In the expression 3.27 if we take derivative with respect to \( \beta \), we get the following:

\[
(1 - \delta)(1 - 2\beta) \leq \beta^2 \delta (1 - \delta^{k-1})
\]  \hspace{1cm} (3.28)
in order to get the desired result. \( \forall \beta \geq 0.5 \), the result is trivial. Otherwise, 3.28 should be satisfied for the desired result. The rationale behind this is that when the agents with more severe self-control problems face with the same project, since their immediate gratification tendency is higher, in order to induce them to complete the investments, the principal should give more immediate rewards to them -which is the bonus here.

### 3.5 Discussion and Conclusion

Our main purpose in this paper is to investigate the role of different preference structures -other than classical time consistent preferences- in bargaining and investment games. We use the well-known quasi-hyperbolic discounting function to incorporate the time inconsistency into our framework. By keeping the environment as simple and general as possible, we explore the behavioral characteristics of different economic agents when they face with intertemporal decisions and with bargaining situations.

We believe that introducing time inconsistency in a strategic environment is the most important aspect of this paper. Incorporating boundedly rational agents into this kind of a game provides us observational differences among them. In case of no bonus, while exponential and naive hyperbolic agent finishes the investment game without any delay, sophisticated agent has a periodic investment plan -explained in section 1.5 and in Theorem 1. When there is bonus scheme, an increasing reward or bonus is necessary for NHA to continue investing. On the other hand, Since exponential agent is time consistent, a fixed amount of bonus makes her to finish the project. The behavior of SHA will be examined later.

The degree of time inconsistency factor is important in determining the behavioral differences. Depending on the severity of the self-control problem, optimal incentives schemes and behavior of the agents change. In the bargaining game, we apply a different kind of equilibrium concept because of bounded rationality. Beliefs of naive hyperbolic agents turn out to be wrong but again by using best response argument, we find the subgame perfect equilibrium. An interesting observational difference about the naive hyperbolic agent is that she is mistaken in predicting her wage. She overestimates it and because of this, she is disappointed
about her realized wage. This misperception leads to a regret motive that she may pursue a goal that is not worthwhile to pursue, since she will get a less payoff than she expects. Thus, this observation may be helpful in understanding people’s disappointments resulting from their great expectations about the future.

The puzzling question in this context is that why does not NHA learn from her experiences? She always behaves consistently in being time inconsistent. It is not that she is not learning but she is always defeated by her tendency to pursue her immediate gratification. She knows and remembers what happened in the past but since she is highly overoptimistic about herself in being time consistent in the future, she basically ignores her past actions and does not take lessons from her previous experiences. Introducing partial naivete may be a more realistic approach in order to incorporate the learning or bounded memory motive.

There are several possible extensions that can be made in our framework. Some of them are the following: Incomplete information about types can be introduced into the bargaining stage. Different cost structures (endogenous, stochastic...etc) could be examined. Underestimation or overestimation of the costs depending on the agents’ types can be examined. Partial naivete can also be introduced into the model.

3.6 References


Periods that for SHA, it is optimal to invest the last (k-1)th unit.

Figure 3.1. Optimal investment schedule
Vita

Zafer Akin

Zafer Akin was born on August 12, 1978, in Istanbul, Turkey. He earned his B.S. degree in Chemistry in 1999 from Bilkent University, Ankara, Turkey. After graduation, he enrolled as a graduate student at the same university and earned his M.A. degree in Economics in 2001. Since then, he has continued his studies at Penn State University as a graduate student. His Ph.D. thesis has focused on time inconsistency in game theoretic frameworks.