Imperfect Information Processing in Sequential Bargaining Games with Present Biased Preferences

Zafer Akın

Department of Economics
TOBB University of Economics and Technology

Working Paper No: 08-10
TOBB University of Economics and Technology
Department of Economics

August 2008

Imperfect Information Processing in Sequential Bargaining Games with Present Biased Preferences

Zafer Akin*

TOBB University of Economics and Technology, Department of Economics, Sogutozu, Ankara, 06560, Turkey

August 14, 2008

Abstract

This paper studies an alternating-offers bargaining game between a time-consistent and a time-inconsistent player who processes information on own preference type imperfectly that flows during the course of the game. The time-inconsistency is modeled by quasi-hyperbolic discounting and the "naive backwards induction" is used as the solution concept. By using cognitive and mood state approaches, we model learning of naive time-inconsistent player. The main result specifies a clear connection between the model parameters and the equilibrium in the most general sense. For special cases, critical delay times and probabilistic structure of the agreement and relations between them are characterized. Comparative statics imply that more naive players and players with milder self control problems are offered higher shares more frequently.

JEL Classification: C78, D83

PsycINFO Classification: 2340, 2343

Keywords: Quasi-hyperbolic discounting, imperfect information processing, sequential bargaining

1 Introduction

Are the standard preferences able to explain people's observed choices? Do people have systematic biases in their decision making? Do they learn about these biases that can be potentially taken advantage by other parties in strategic environments? How does this affect the equilibrium outcome? Does one bias exacerbate or alleviate another bias' good or bad effects? Except the first two, whose answers are provided by a growing body of both anecdotal and experimental evidence supporting deviations from standard preferences, these questions are not addressed enough in economics literature. The aim of this paper is to answer these

*E-mail: zafer.akin@etu.edu.tr

comprehensive questions partially in a strategic environment where agents, who potentially have present biased preferences and process information imperfectly, play a sequential bargaining game.

There are many systematic self-serving biases mentioned both in psychology and economics literature one of which is the present biased preferences. It is frequently observed in experimental studies that there is a sharp short run drop in valuation and a faster decline rate in the short run than in the long run. In other words, people have preference for immediate gratification and they tend to pursue it although they may regret about this pursuit in the future. Theoretically, agents with present biased preferences have a relative preference for payoffs (benefits and costs) at an earlier period over payoffs at a later period with one important exception that the relative weight given to the earlier period's payoff increases as the earlier period gets closer. The observations mentioned above are properly captured by quasi-hyperbolic discounting model, which is frequently used way of modeling the present biased behavior. This kind of behavior can be seen as a lack of self-control and leads to time-inconsistency.[1] There is a growing literature incorporating this type of preferences in a wide range of contexts and some fruitful explanations about seemingly anomalous behavior that are not fully explained by standard economic theory have been produced (Strotz, 1956; Laibson, 1997; O'Donoghue and Rabin, 1999a).

One approach in this literature synthesizes quasi-hyperbolic discounting and noncooperative games. Since individuals suffer from present biasedness not only in their personal decisions but also in strategic environments, this approach is an interesting and a valid application. There is a very new but growing literature on noncooperative games played by agents who have present biased preferences. Some examples are as follows: Akin (2004) incorporates quasi-hyperbolic discounting into a self-investment game followed by a sequential bargaining game; Dellavigna and Malmendier (2004) analyze the contract design of firms when they face time inconsistent consumers; Sarafidis (2005) examines the intertemporal price discrimination game between a durable good monopolist and time inconsistent consumers; Akin (2007) introduces learning in a sequential bargaining framework and shows that learning is a possible explanation for bargaining delays and Chade et.al. (2008) study infinitely repeated games with perfect monitoring, where players have quasi hyperbolic discounting.

In the quasi-hyperbolic discounting framework, in addition to time consistent exponential type, which is the benchmark case, three types of agents with time inconsistent preferences are considered: sophisticated ones who are aware of their time inconsistency, naive ones who are completely unaware, and partially naive ones who are partially aware[2] of it (O'Donoghue and Rabin, 2001). Akin (2004, 2007) and Sarafidis (2005)

[1] Agents with time inconsistent preferences, quasi hyperbolic discounting, self control problems are used interchangeably in the paper.

[2] Being completely aware of self control problems is actually not a systematic bias since the agent can rationally make decisions by taking into account this own characteristic. However, being naive or partially naive about these problems is a systematic bias in the sense that the agent can only make rational decisions for the current self but not for the future selves

assumed all types except partial naives but Dellavigna and Malmendier (2004) examined especially partially naive types. We consider all types in our framework.

Regarding the mentioned questions at the beginning, some learning framework has to be incorporated into the time inconsistent preferences. Therefore, we consider a dynamic structure of the preferences by incorporating learning of partially naive players about their own type. In the existing models, there are a few papers that just mention learning and its effects on the equilibrium behavior (O'Donoghue and Rabin, 2001, Dellavigna and Malmendier, 2006). Learning of hyperbolic discounters is first examined in a noncooperative game context by Akin (2007). He introduced naive learning in a bargaining framework and characterized the conditions for bargaining delay caused by naive learning where the naive agent systematically learn her own type during the game.

The hyperbolic agent presumably has a given time preference structure and agents are categorized into different types based on their perceptions on this structure. The literature takes these types as given and examines implications of these. On the other hand, incorporating learning is a dynamic and a more reasonable approach. This paper attempts to improve the dynamic theory at hand by adding another behavioral dimension. It not only makes the perceptional dimension more sophisticated but it also incorporates a psychological characteristic most people have. Specifically, we assume that there are two possible (cognitive) states of the world. In one of the states, the agent is hyperbolic, in the other; she is either exponential or hyperbolic. In addition, independent of her cognitive state, she can be in two different psychological states (mood states). At each mood state, she potentially process information imperfectly and differently.[3]

Mood state approach can be interpreted in two different ways. First, the agent who has two different psychological moods at each time may or may not take into account a specific/relevant information depending on her mood in her decision making process. At each of her moods, she disregards (treats) the information (as irrelevant) with potentially different fixed probabilities. Second, the agent tends to misinterpret some information contradicting her prior beliefs based on the mood state ("disconfirmation bias", Wason, 1960). In other words, she may misinterpret the available information, probably as supporting previously held beliefs, with different probabilities depending on her mood (e.g., she has a more severe tendency of misinterpretation in one mood state than she has in the other).

The fine distinction between two interpretations is that in the former, the agent either accepts the information or rejects it. When she accepts, information is processed as it should be and perfectly, and

---

and, without learning, this occurs systematically.

[3] We assume that mood states follow a Bernoulli process, a sequence of independent identically distributed Bernoulli trials (States themselves can be correlated and state transitions can be defined accordingly but we will not pursue this approach in this paper .See Compte and Postlewaite, 2008, for a model that incorporates this approach). We also assume that there is no correlation between the cognitive states and the mood states. As an extension, a framework with correlated cognitive and mood states can be analyzed (e.g., a more naive agent may tend to be in a mood state where she ignores/misinterpret the information with a higher probability).

accepting or rejecting the information is probabilistic.[4] In the latter, the agent accepts the information for sure but she may misinterpret it. She correctly and perfectly processes the information if it supports previously held beliefs. However, she may misinterpret the information and process it as supporting her prior beliefs even if it does not actually support them. Here, whether the agent misinterprets the information is probabilistic.

Given these behavioral assumptions, this paper examines the behavior of agents having, potentially, self control problems (along with the mentioned characteristics) in a bargaining framework. We focus on alternating-offers bargaining game as the strategic interaction environment for the following reasons. It is a sequential (not a one shot) game played multi periods (potentially, infinite horizon) and this allows the modeler to observe the effect of present biased preferences explicitly. It is a two person game where different types of hyperbolic agents can interact (not a one person game played by different selves of one player). It is a dynamic game in the sense that during the course of the game, feedback and information is generated about the underlying uncertainty if there is any. Hence, it allows learning. Since it has a wide range of real life applications, there is a large literature about sequential bargaining games and it has a well established theory.

In the present bargaining model, two players -one is partially naive, the agent, and the other is exponential, the principal- are engaged in an infinite horizon alternating-offers bargaining game. The agent is partially naive due to the way how her cognitive system works. Particularly, she has two cognitive states. In one state, she is hyperbolic (hyperbolic state in which she has the same beliefs with a sophisticated agent). In the other, she is either exponential or hyperbolic (mixed state in which she has partially naive beliefs in which she assigns a lower probability to being hyperbolic than it should be). She becomes more sophisticated over time by observing possible rejections. Specifically, the agent has prior beliefs on the likelihood of the cognitive states and she updates this prior[5] with the feedback flowing during the course of the game, which is filtered depending on the mood state.

We assume that each player knows other players' available strategies and their current preferences. However, beliefs about their own and others' future preferences may differ across different types. Exponential and sophisticated players' current and future preferences are common knowledge among all players. On the other hand, naive and partially naive players misperceive their own future preferences and believe that their beliefs about themselves are shared by all types of players. The beliefs about the mood states are such that the exponential player knows everything that the modeler knows but the partially naive player inherently

[4] If the ignorance of the relevant information is interpreted as mistakes by the agent, then we can make the following argument: rather than a gradual learning, there may be a threshold accumulated information (disregarded/ignored since they are treated as mistakes) level above which the agent inevitably concedes the implication of the accumulated information.

[5] Except the initial period, potential learning can be attributed to the rejections of each party. She may learn not only when her offer is rejected but also when she rejects an offer.

has these characteristics and is not aware of them.

Since the stationarity of preferences is violated in the framework of time inconsistency, the choice of the solution concept is tricky. In this paper, we use "naive backwards induction" (NBI) as the solution concept developed by Sarafidis (2005) and extended by Akin (2007) for games played by time-inconsistent agents. In NBI, players play best response to what they think the opponents will play and the strategies have to be sequentially rational.

By using NBI and the results from Akin (2007)[6], we find the following results. In the most general case where imperfect information processing is allowed, there is a critical date before which the bargaining game does not end for sure but this critical value may well be zero depending on the parameter values. Moreover, after this threshold date, the agreement is probabilistic and the average delay time is calculated accordingly. The critical value for the learning probability parameter is characterized so as for the game to end immediately. When we examine the perfect Bayesian case, we have a similar result as in Akin (2007) and show that this critical date is always larger than or equal to the one found in the imperfect information processing case. However, the actual waiting time may well be greater in the imperfect processing case than the one in the perfect processing case. In sudden learning case, we find that for sufficiently small learning probability, there is immediate agreement. If this probability is larger than a specified threshold level, then the agreement is probabilistically delayed. Comparative statics imply that against naive agents, the exponential agent offer a higher share more frequently (we mean "for a wider range of learning probability" by "more frequently") to the more naive ones (higher $\widehat{\beta}$) and to the ones with less severe self control problems (higher β).

The remainder of the paper is organized as follows. In Section 2, we summarize the related literature both in psychology and economics. In Section 3, we introduce the model in detail. In Section 4, we first characterize the equilibrium behavior in the most general case. We then examine the perfect Bayesian case where there is perfect information processing. In the last part, we study a special case where the cognitive states are mutually exclusive that leads to a phenomenon called "sudden learning" and some comparative analysis are presented. Section 5 discusses the implications and limitations of the presented model and some extensions.

2 Related Literature

People inherently have self-serving biases, which constitute a subset of behavioral biases, as opposed to the perfect rationality assumption of the standard economic theory. Obviously, given these biases, agents who

[6] The solution concept NBI is especially used in finding the perfect equilibrium partitions in bargaining (Akin, 2007, has all the details.) The rest here is basically intertemporal optimization given the stochastic structure of preferences.

are aware of them can act rationally. However, the very nature of the existence of these biases is to be unaware of them in most of the cases. Among these, self-serving biases are the most inborn ones that are defined as holding beliefs or having directed preferences that favor one's own payoff or satisfaction. Biased beliefs may be about the payoffs, other players' types, available strategies, environment and many other things. Present biased preferences, disconfirmation bias and overconfidence are some well known examples.

Self-serving beliefs influence one's judgments in many different ways. Present bias affects the relative weight of payoffs at different times that results in a judgment fallacy making people to discount near future payoffs more heavily than the more distant ones. Overconfidence refers to a self-serving belief that makes people overestimate their real performance. For example, overconfidence makes people to think that they have better skills or knowledge than they actually have (Myers, 1996), or it can make people to ignore relevant information in a decision making situation (McGraw et. al., 2004). Disconfirmation bias makes people to disregard some information contradicting their previously held beliefs (Wason, 1960) and when people base their decisions on these biased beliefs, their judgments would naturally be biased too. There are no clear cut distinctions among these biases and most of the time, their implications overlap.

One of the implications of self-serving biases we focus on is imperfect information processing.[7] Closely related to the ones above, we now mention some other considerations in psychology literature referring to limited/imperfect information processing.

People's psychological states can affect information processing, thereby effort and performance. Nonpermanence of failure or permanence of success phenomenon refers to attributing the failures to things that are less likely to occur again in the future and making no such attributions in successes. Alternatively, pervasiveness of success bias refers to attributing successes to a broader set of circumstances than is appropriate (Seligman, 1990). These can be viewed as misinterpretation of available information.

In Benabou and Tirole (2002), a (Bayesian) agent makes a self evaluation on how able he is and he may consciously decide to forget some negative experiences (selective memory). In Koszegi (2000), the agent stops recording any further signal. However, the agent is conscious about how the information is processed and takes this into account when forming beliefs (beliefs are not biased on average).

Negative affective states may change the process how people make decisions. It is usually argued that negative affective states such as anxiety and sadness influence people's ability to process information (e.g.,

[7] There may be plenty of reasons of imperfect information processing (Antonides, 1996). Self-serving biases cover a wide range of those cases. However, here, we will approach imperfect information processing phenomenon from a more affective, visceral and noncognitive perspective. Present bias as a self-serving bias may itself cause imperfect information processing but we discriminate these two and assume no mutual interaction between them. We treat the present bias as a cognitive process and assume that it is not directly influenced and driven by the mood or affective conditions (states). On the other hand, we treat imperfect information processing as a noncognitive and affective process that is not directly affected or driven by present biasedness. For a more formal discussion of information processing imperfections as the source of bounded rationality, see Lipman (1995) in which he discusses partitional models, non-partitional models, and axiomatic approaches.

Ellis and Ashbrook, 1988; Raghunathan and Pham, 1999). For example, there is some evidence that anxious or sad individuals process information less systematically in judgment and decision making (e.g., .Sanbonmatsuand and Kardes, 1988; Schwarz, Bless and Bohner, 1991).

According to Schwarz and Clore (1983), people routinely use their feelings as information and beliefs affect one's feelings, thereby indirectly affecting judgments. On the other hand, in mood congruent judgments, mood congruent beliefs are assumed to be the real bases of judgments (Schwarz and Clore, 2003). Maia and McCleland (2004) have parallel findings in which they ask participants for their feelings about the past twenty trials in IOWA gambling task. They show that when participants asked about their feelings on choices, they use their feelings when they decide among alternative options. In our interpretation of feeling as information theory, people sometimes interpret the relevancy of the affective information differently according to their current mood condition. Thus, one's mood condition indirectly affects that person's judgments by first affecting the situation's affective information value. For example, a person can misinterpret her feelings about her choices because her current mood is contradicting with her feelings on choices. Moreover, there is a wide amount of empirical data showing that people treat alternative options differently when they are induced different mood conditions (e.g., in a positive mood condition, they become more risk averse while they become more risk seeker in a sad mood condition, Isen, 2002). In this paper, we make information processing dependent on the mood state in a stochastic way such that if someone is in a current mood state, there is a fixed probability that she correctly interprets her feelings on choices.

The other issue we focus on is, so called, cognitive states that are related to the perceptions on the future preferences. It is generally considered that agents' decisions are purely based on "cold" cognitive processes but this may not always be the case. Most decisions elicit feelings before and after the decision making process starts. This would happen in a variety of ways (Schwarz 2000). For example, the expectation for a desired outcome can produce positive affect towards that choice before we decide on that option or a decision can induce negative mood if our expectations do not coincide with the results of our choices. We could claim that emotions are somehow interacting with our cognitive system. The interplay between "emotional brain" and "cognitive brain" is best handled by the dual self or dual system models[8] (Metcalfe and Michel, 1999; Kahneman, 2003; McClure et al., 2004).

[8] In our context, another way of modeling that allows interaction between emotional brain and cognitive brain can be summarized informally as follows: a partially naive agent may act as a sophisticated agent (or become more aware of her type by updating her beliefs) when her mood contradicts with her feelings about choices. Even if a partially naive agent currently believes that she will not procrastinate in the future that is an optimistic belief, now, she may not procrastinate since she now might be actually sad about her self-control problem. In other words, when her affective interpretation about her feelings on choices contradicts with her mood state, she can interpret her beliefs more rational than usual. Otherwise, she can underestimate her self-control problem because she holds a positive belief in which she thinks she have control on it.

The model we propose in this paper only allows an indirect interaction between these two systems, not a direct one as mentioned above.

Metcalfe and Michel (1999) propose that there are two systems in the brain, a cool cognitive "know" system and a hot emotional "go" system where the interaction between these two systems play a critical role in self-control behavior in intertemporal decisions. The general characteristics of those systems are such that the former is more complex, slow, develops late and controls the self-control behavior while the latter is simple, reflexive, fast, developed early, and accentuated by stress. Their dual system model explains very fundamental processes underlying self control behavior and similar models have been recently utilized in economics to model addictive behavior and explain observed systematic biases.[9]

In a related model, Lowenstein et. al. (2003) investigates the qualitative nature of change of future preferences of people who exaggerate the degree to which their current tastes will represent their future tastes. They argue that people overestimate the magnitude of this representation, which leads to "projection bias". In their model, They use state dependent utility functions (different states can be interpreted as hot and cool states) and projection bias means that predicted utility lies in between the utility given the current state and the true future utility. They find that this tendency leads people to consume too early in life and leads to improper purchasing of durable goods. Our model is similar to this phenomena in the sense that although the present bias will always emerge, the agent optimistically thinks the current present bias will vanish in the future. In other words, the agent incorrectly thinks that her current state will change in the future but actually it will not. In their context, Lowenstein et. al. (2003) emphasize that learning with experience is not likely to occur although people are aware of this bias in a meta level.[10]

Finally, Neale and Bazerman (1985) examine experimentally the effect of framing and overconfidence on bargaining behavior and outcomes and find that overconfidents exhibit less concessionary behavior and perform worse than realistically confidents do. In a related paper, Babcock and Loewenstein (1997) employ self-serving biases to explain bargaining impasse. They provide experimental evidence showing that self-serving biases cause impasse and result from selective evaluation of information, together implying that information processing in a self-serving manner causes bargaining delay or disagreement. Similarly but theoretically, we, in this paper, show that an agent's biased information processing in the direction of her own self-interest[11] causes delay in agreement in a sequential bargaining framework.

[9] Bernheim and Rangel (2004) consider an addiction model where the agent is sometimes in a hot mood in which she consumes the addictive good as opposed to her cool self's will. Fudenberg and Levine (2006) proposes a simple dual-self model giving a unified explanation for several observed irregularities, such as time-inconsistency and paradox of risk aversion in the large and small rewards.

[10] By learning, Lowenstein et. al. (2003) mean the bias to diminish overtime but we use learning as becoming aware of the underlying bias. Learning in their sense can also be considered in our context but we do not believe that the present bias does diminish overtime with experience. In other words, we argue that β is fixed meaning that there is no learning, but that $\widehat{\beta}$ converges to β overtime with experience meaning that there is increase in awareness.

[11] Perfect information processing refers to an unbiased evaluation of information but biased information processing is self-serving in our context because, the information flowing during the game will always indicate that the agent actually is optimistic about her future preferences. Since the level of sophistication is negatively correlated with the equilibrium share, this bias is in

On top of these linked and growing branches of the literature, the part of our model that deals with the perceptions on future preferences proposes that the agent, cognitively, can be in two different states (founded on these perceptions) at any given moment; hyperbolic state (cool state) and exponential state (hot state).[12] Moreover, there are two states of the world; pure hyperbolic state and the mixed state.[13] The other part of our model posits that the agent imperfectly processes the flowing information based on the mood state and then uses this information cognitively to update her beliefs on what the actual state of the world is. Given the actual state of the world is the hyperbolic state, the process on the cognitive side during the game can be called as "cooling down" (learning her actual preference structure), which means that the agent becomes more sophisticated over time.

Generically, we use "learning" to actually stand for increasing self awareness. In our context, the agent's self awareness about what her actual future preferences for immediate gratification are increases overtime. In an environment with an overconfident agent, learning means the improvement of the agent's self awareness about her ability or her potential to achieve something. For an agent suffering from disconfirmation bias, learning denotes the enhancement of the agent's self awareness about her judgmental imperfections leading to more rational decision making.

3 Model

Let $T = \{0, 1, 2, ...\}$ be the infinite set of possible agreement times. Let $i \neq j \in \{1, 2\}$ and $t, s \in T$ represent players and dates, respectively. Let $u = \left\{u^1, u^2\right\} \in U$ be a utility pair and U be the set of feasible utility pairs where $U = \left\{u \in [0,1]^2 \,|u^1 + u^2 \leq 1\right\}$. At each $t \in T$, i offers a utility pair. If j accepts the offer, the game ends. In case of rejection, j offers a utility pair at $t + 1$. Each player gets nothing if there is no agreement.

We assume two different types of players.[14] One player is partially naive hyperbolic (PNHA), who can potentially become sophisticated during the game, and the other is exponential player (EA). The PNHA is not fully aware of her[15] time-inconsistency. She thinks that she will be impatient in future periods but she underestimates its level. The sophisticated hyperbolic agent (SHA) is fully aware of her time-inconsistency and

the direction of the agent's own self interest.

[12] Hyperbolic state refers to the notion of being sophisticated and exponential state refers to the notion of being naive. We call the exponential state as the hot state because the agent is naive and totally optimistic about her future preferences in this state. In the hyperbolic state, the agent is sophisticated and totally aware of her future preferences that is why we call it as the cool state.

[13] The mixed state (with some probability, she will be in exponential state in the future, with the complement probability, she will be in hyperbolic state) can be interpreted as the agent being partially naive.

[14] There are actually four types: exponential, naive, partially naive and sophisticated agents. Naive and sophisticates are basically the extreme cases of partially naive types. That is why we are only interested in partially naive ones.

[15] We call the partially naive and sophisticated agent as "she" and the exponential agent as "he" for convenience.

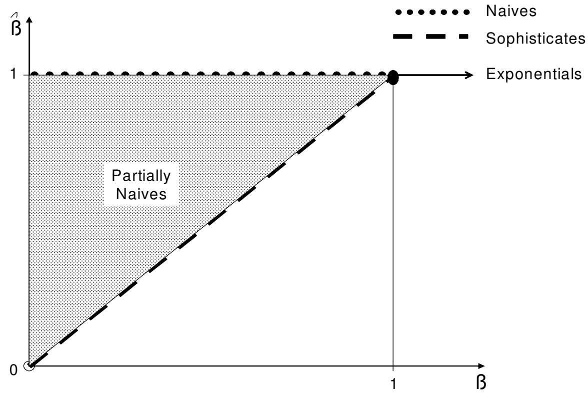

Figure 1:

behaves accordingly. The EA is time-consistent and has the discount factor sequence $\{1, \delta, \delta^2, \delta^3, ...\}$ where δ is the standard time-consistent impatience. The PNHA and the SHA has $\{1, \beta\delta, \beta\delta^2, \beta\delta^3, ...\}$ where β is the time-inconsistent preference for immediate gratification or the self-control problem of the agent. $\widehat{\beta}$ (degree of naivete) represents an agent's belief about what her β will be in all future periods. Thus, the EA has $\beta = \widehat{\beta} = 1$, the SHA has $\beta = \widehat{\beta} < 1$ and the PNHA has $\widehat{\beta} \in (\beta, 1)$. While the SHA is located at one extreme of the scale of awareness on the future preferences, the naive agent (NHA) is at the other extreme of this scale where she is completely unaware of her time-inconsistency, $\beta < \widehat{\beta} = 1$. Figure 1 is the visual representation of the actual and perceived self control problems.

We assume that each agent knows other players' available strategies and their current preferences. However, beliefs about their own and others' future preferences may differ across different types. We now briefly summarize these beliefs. Naive agents believe that they will not have time inconsistency in the future and that this is common knowledge at every information set. Partially naive agents believe that it is common knowledge that they will have some self-control problem in the future. NHA and PNHA recognize time inconsistency of the others, if any, but not their own and they know all the opponents' future preferences correctly. Sophisticated players certainly believe at every information set that it is common knowledge that they will have time inconsistency in the future. The previous statement is common knowledge among all EA and SHA. In addition, EA is more informed than the naive player in the sense that he knows the naive opponent's learning potential but the naive agent herself is not aware of this.[16] The psychological state

[16] Over-optimism or over-confidence or both may cause a person to be naive about self-behavior in the future. It is agreed that "on nearly any dimension that is both subjective and socially desirable, most people see themselves as better than average" (Myers, 1996). Since patience can be seen as a kind of ability, it is plausible to think that people may consider themselves more patient than others. [See Brocas and Carrillo (2004), Zabojnik (2004) and Pinto and Sobel (2005) for a more detailed

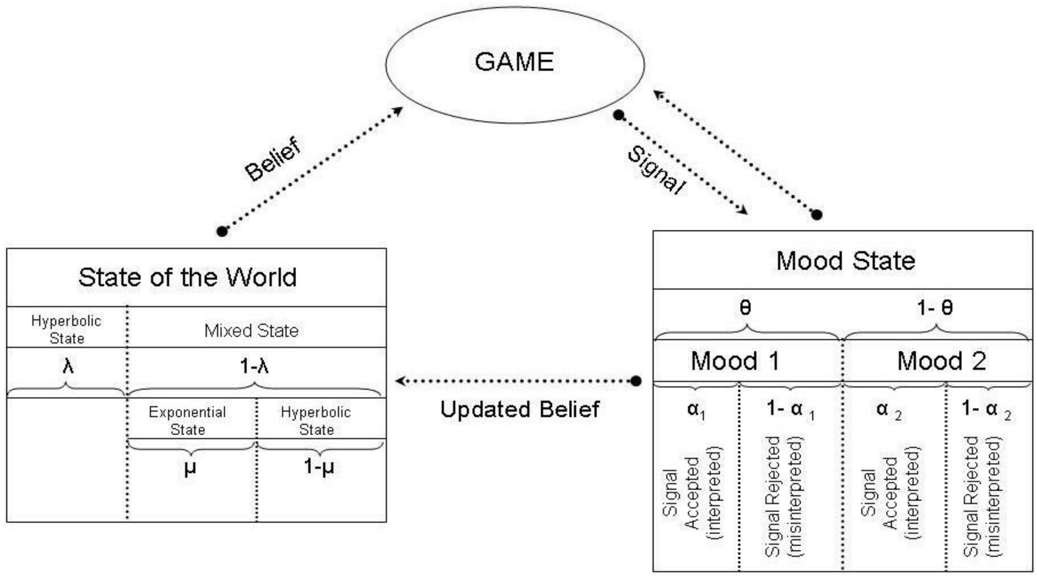

Figure 2:

structure of the partial naive agents is not common knowledge among agents such that the partially naive player is not conscious about them but the exponential player knows everything that the modeler knows.

Naive backwards induction is used as the solution concept. In NBI, each agent plays a best response to what he/she thinks to be the opponent's play and the strategies are required to be sequentially rational. One important point is that due to self-misconceptions and misperceptions about the opponent, naive agents may not be able to anticipate the opponent's actions correctly and they may be surprised during the course of the game. Akin (2007) elaborates on both the solution concept and beliefs in detail.

Figure 2 summarizes the model. There are two possible states of the world. One is hyperbolic state where $\widehat{\beta} = \beta$. The other is the mixed state where $\widehat{\beta} = 1$ with probability μ and $\widehat{\beta} = \beta$ with probability $1 - \mu$. The agent's prior belief on the two states of the world is λ_0 and $1 - \lambda_0$, respectively. The *actual* state of the world is the hyperbolic state where the PNHA will always employ β as the preference for immediate gratification as opposed to her current beliefs, $0 < \lambda, \mu < 1$. In addition, the actual state of the world is known by the EA. We allow belief updating based on the signals received during the course of the game both in Bayesian sense and a boundedly rational sense.

The signals received during the game are processed by the PNHA depending on the mood state. There are two possible moods of the agent whose respective probabilities are θ and $1 - \theta$. In different moods, she

discussion on over optimism due to positive self-image and related discussions.]

accepts (interprets) the signal (correctly) with potentially different probabilities α_1 and α_2, respectively. With complement probabilities, the agent disregards (interprets) the signal (incorrectly).[17] Without loss of generality, we impose $\alpha_1 \geq \alpha_2$. The α_i values being strictly less than one embeds a bounded rationality aspect into the agent's behavior.[18] Since, by construction, the information flowing during the game always signals the agent's disadvantageousness, α_i being strictly less than one implies a self-serving assessment in bargaining.

Note that $0 \leq \lambda, \mu, \theta, \alpha_i \leq 1$, $i = 1, 2$. The θ and α_i parameters are related to the behavioral characteristics of the agent, exogenously given and do not change overtime.[19] The λ and μ parameters are related to the beliefs of the agent. The probability μ is exogenously given and do not change overtime but the probability λ is the belief of the agent regarding the states of the world and it is subject to change based on the signal processing.

Underlying uncertainty (state of the world) in the model is all about the PNHA and her perceptions. It is common knowledge that the exponential player is exponential, and that he does not have any behavioral characteristics like the PNHA. In other words, the exponential agent does not have any mood state ($\theta = 0$ or $\theta = 1$) and he perfectly processes available information ($\alpha_1 = \alpha_2 = 1$). Moreover, the EA knows the actual state of the world. In addition, the followings are imposed regarding the information structure: The characteristic of the EA and the probability assessments of the PNHA on the states of the world are common knowledge. The behavioral characteristic of the PNHA is known by the EA. Moreover, at each even t, the PNHA thinks that the game will end with the terms arising from the realized preferences at t but she may be surprised. At each odd t, she offers based on the realized preferences at t.

[17] By accepting the signal or interpreting it correctly, we mean that the agent updates her beliefs towards the true state (hyperbolic state). By disregarding the signal or misinterpreting it, we mean that the beliefs do not change. Thus, in a sufficiently long time, she comes to believe in the true state almost surely (the game will not be delayed infinitely, after sufficient learning periods, it will end). This is because of the structure of the game in which each period the game is played makes the agent update her beliefs towards the true state. In a different game setting, there may be two or more conflicting signals that result in ambiguous dynamics of the beliefs (Rabin and Schrag, 1999).

[18] Here, we focus on a sequential bargaining game and the outcomes of the game (offers and accept/reject decisions) turn out to be the signals about the states of the world. Given the assumption of an informed principal and the inherent characteristic of the agent, the outcomes always signal the true state of the world (t period delay in agreement means t rejections by the PNHA herself and the EA. Any confirmation of the agent's beliefs ends the game).

[19] Given this exogeneity, we implicitly assume that the signal and the mood state probabilities are not correlated. In a more sophisticated and realistic model, this assumption can be relaxed in such a way that the received signal and the likelihood of different mood states may be correlated.

4 Equilibrium Characterization

In this part, we first keep everything as general as possible and impose no restriction on the parameter values, $1 > \lambda, \mu, \theta, \alpha_i > 0$, $i = 1, 2$. This allows a boundedly rational approach where the agent may ignore or disregard available information depending on her mood state. We then look at a case where the PNHA is a perfect Bayesian learner in the sense that she updates her beliefs on the states of the world based on the information flowing that is perfectly processed (e.g., $\alpha_1 = \alpha_2 = 1$). Then, we examine a special case where the PNHA learns her own characteristic with a fixed probability (sudden learning) at each period. Finally, we report some comparative analysis results.

Before proceeding, it will be useful to mention the equilibrium shares of the Rubinstein's alternating-offers bargaining game (1982) where players are exponential and have discount factors δ_1 and δ_2 and then the results from Akin (2007) when the players are potentially hyperbolic discounters:

Remark 1 *In the infinite horizon alternating-offers bargaining game where both players have exponential discounting with discount factors δ_1 and δ_2, the equilibrium payoffs are $(x^*, 1-x^*) = (\frac{1-\delta_2}{1-\delta_1\delta_2}, 1-\frac{1-\delta_2}{1-\delta_1\delta_2})$. The bargaining game between an EA and a PNHA ends always with shares$(x^*, 1-x^*) = (1-\beta\delta\frac{1-\delta}{1-\widehat{\beta}\delta^2}, \beta\delta\frac{1-\delta}{1-\widehat{\beta}\delta^2})$ if the EA offers first; if the PNHA offers first, she offers $(x^*, 1-x^*) = (\frac{1-\delta}{1-\widehat{\beta}\delta^2}, \frac{\delta-\widehat{\beta}\delta^2}{1-\widehat{\beta}\delta^2})$ but the EA rejects this offer, where x^* is the share of the first proposer. Moreover, a bargaining game between an EA and an SHA ends with the following shares $(x^*, 1-x^*) = (\frac{1-\beta\delta}{1-\beta\delta^2}, \frac{\beta\delta-\beta\delta^2}{1-\beta\delta^2})$ if the EA offers first, $(x^*, 1-x^*) = (\frac{1-\delta}{1-\beta\delta^2}, \frac{\delta-\beta\delta^2}{1-\beta\delta^2})$ if the SHA offers first.*

From here on, *all mentioned shares are of the EA who is the first proposer.* For notational convenience, we denote $\overline{S} = \frac{1-\beta\delta}{1-\beta\delta^2}$, $\underline{S} = 1-\beta\delta\frac{1-\delta}{1-\widehat{\beta}\delta^2}$ and $\widetilde{S} = \frac{\delta-\widehat{\beta}\delta^2}{1-\widehat{\beta}\delta^2}$. $\overline{S}$, $\underline{S}$ and $\widetilde{S}$ are all the EA's share when he plays with an SHA and is the first proposer; he plays with a PNHA and is the first proposer and he plays with a PNHA and is the second proposer, respectively. The following relationships are satisfied $\overline{S} > \underline{S} > \delta\overline{S} > \delta\underline{S} > \widetilde{S}$.

4.1 The Generic Case

In this subsection, we do not impose any restrictions on the model parameters and treat them as variables in the most generic sense. Specifically, we do not allow any parameter to take an extreme value, 0 or 1. This means that the cognitive states are not mutually exclusive (they are overlapping) and the agent's mood states can vary across and within the game and across periods.

The agent processes information in a limited sense depending on the mood state. She accepts (interprets) the information (correctly) with fixed and potentially different probabilities at each mood state, α_1 and α_2. In which mood state she is at each period is also probabilistic, θ and $1-\theta$. This indicates that she processes the information with probability $\rho = \theta\alpha_1 + (1-\theta)\alpha_2$.

The belief on the state of the world is initially given by the probabilities λ_0 and $1-\lambda_0$. At each period that the agent processes the information, she updates her belief on λ such that if until period t, the agent has processed the information $s \leq t$ times, then the updated belief is given by $\lambda_s = \frac{\lambda_0}{\lambda_0+(1-\lambda_0)\mu^s}$. Given this updated belief, the effective discount factor of the agent is given as follows:

$$\begin{aligned}\delta_s &= \lambda_s\beta\delta + (1-\lambda_s)(\mu\delta + (1-\mu)\beta\delta) \\ &= \beta\delta + \mu\delta(1-\beta)(1-\lambda_s)\end{aligned}$$

The following theorem states that there exists a date before which there is no agreement. This threshold date may well be zero. The logic behind this result is the trade off that the principal faces. By delaying, he can probabilistically make the agent more sophisticated whom he can extract more share. However, the effective discount rate is smaller than one so that the overall size of the pie is shrinking. Thus, this trade off defines a threshold date (not probabilistic). The theorem also states that after this threshold, the agreement date is probabilistic and this probability is characterized. This result is similar to the main theorem in Akin (2007) with the exception that there is no probabilistic agreement after the threshold in the latter.

Theorem 1 In the game played between an exponential player and a partially naive player (having behavioral characteristics specified above), in any stationary NBI solution, there exists an integer $t^*_{GC} \geq 0$ such that the game ends at period $t \geq t^*_{GC}$, $t \in t^*_{GC} + 2i$, $i = 0, 1, 2...$ if t^*_{GC} is even ($t \geq t^*_{GC} + 1$, $t \in t^*_{GC} + 2i + 1$, if t^*_{GC} is odd) with probability $f(t)$, $t\tilde{\ } NegBin(t^*_{GC},\rho)$ where $f(t)$ is the probability mass function and $NegBin(t^*_{GC},\rho)$ is the negative binomial distribution with parameters t^*_{GC} and ρ, and there needs to be $\frac{t^*_{GC}}{\rho}$ periods for the game to end on average. If $t^*_{GC} = 0$, the game ends immediately.

Proof. Let the EA be the first proposer. He offers a share at each even period if the game did not end by then. The other case is very easy to adopt. Note that by lemma 1 in Akin (2007), the PNHA's offers are never accepted. That is, the game never ends at odd periods in which the PNHA makes an offer. Define (s,t) as the state of the game played at period t where the PNHA has processed the information $s \leq t$ times. Then, the EA's problem can be expressed by the following value function:

$$V(s,t) = \max\{R(s,t), \delta^2 E(V|s,t)\}$$

$R(s,t)$ is the immediate reward function that is the best payoff the EA can get by finishing the game at (s,t) and $R(s,t) = x^s_t$ where $x^s_t = 1 - \beta\delta\frac{1-\delta}{1-\delta\delta_s}$ and $\delta_s = \beta\delta + \mu\delta(1-\beta)(1-\lambda_s)$. $E(V|s,t)$ is the expected continuation value of the game and can be written as

$$E(V|s,t) = \rho^2 R(s+2,t+2) + 2\rho(1-\rho)R(s+1,t+2) + (1-\rho)^2 R(s,t+2)$$

$$\delta^2 E(V|s,t) = \delta^2(\rho^2 x^{s+2}_{t+2} + 2\rho(1-\rho)x^{s+1}_{t+2} + (1-\rho)^2 x^s_{t+2}). \tag{1}$$

The following argument shows how this is derived. If the game is delayed one period, then with probability ρ, the EA's expected share will be $x_{t+1}^{s+1} = 1 - \beta\delta\frac{1-\delta}{1-\delta\delta_{s+1}}$ and with probability $1-\rho$, his expected payoff will not change, $x_{t+1}^{s} = x_t^s$. At $t+2$, again with probability ρ, the EA's share will be $x_{t+2}^{s+2} = 1 - \beta\delta\frac{1-\delta}{1-\delta\delta_{s+2}}$ and with probability $1-\rho$, his payoff will stay the same $x_{t+2}^{s+1} = x_{t+1}^{s+1}$ if the PNHA updates her beliefs at $t+1$; if she does not update her beliefs at $t+1$, then with probability ρ, the EA's share will be $x_{t+2}^{s+1} = 1 - \beta\delta\frac{1-\delta}{1-\delta\delta_{s+1}}$ and with probability $1-\rho$, his payoff will stay the same $x_{t+2}^{s} = x_{t+1}^{s} = x_t^s$.

The rationale behind the form of the value function is the following: at each period the EA makes an offer, he will choose either ending the game by making an acceptable offer or delaying the game by two periods by making a rejected offer. The EA can get $R(s,t)$ if he consents to the terms of the PNHA at t. If the EA believes to extract more share from the PNHA in expected terms, he will delay the game. Thus, at time t, the EA will choose the option that gives a higher payoff, x_t^s or expression 1 above. Note that $x_j^i = x_k^l$ if $i = l$ for all j and k, $j \geq i$, $k \geq l$. This implies $x_t^s = x_{t+2}^s$. Thus, the comparison will be between $R(s,t) = x_t^s$ and $\delta^2 E(V|s,t)$ in expression 1. In other words,

$$x_t^s(1-\delta^2(1-\rho)^2) \text{ and } \delta^2\rho(\rho x_{t+2}^{s+2} + 2(1-\rho)x_{t+2}^{s+1}).$$

Remember that the shares are functions of constant model parameters $\beta, \delta, \mu, \lambda_0, \rho$ and number of successful learning periods s. Now, define a function $F(s;\beta,\delta,\mu,\lambda_0,\rho)$ as follows:

$$F(s;\beta,\delta,\mu,\lambda_0,\rho) = x_t^s(1-\delta^2(1-\rho)^2) - \delta^2\rho(\rho x_{t+2}^{s+2} + 2(1-\rho)x_{t+2}^{s+1})$$

This function defines a critical value for s, s^*, that makes the EA indifferent between finishing the game immediately and delaying it. The critical value s^* is defined as

$$s^* = \left\{ \begin{array}{ll} \min\{s \in \{0,1,2...\} | F(s;\beta,\delta,\mu,\lambda_0,\rho) \geq 0\}, & \text{if } F(s;\beta,\delta,\mu,\lambda_0,\rho) \geq 0 \text{ for some } s. \\ 0 & \text{, if } F(s;\beta,\delta,\mu,\lambda_0,\rho) < 0 \text{ for all } s. \end{array} \right\}$$

We now call s^* as t_{GC}^* (GC stands for generic case) that is the number of periods of delay that the EA can tolerate. Given this critical value, the EA is consent to wait until t_{GC}^* periods of successful learning occurs. This means that the EA is ready to delay the game for two periods since the expected discounted value of delaying is larger than finishing now until period t' where $V(t_{GC}^*, t') = R(t_{GC}^*, t') = x_{t'}^{t_{GC}^*}$. Period t' is not known ex ante since it is stochastically determined by the PNHA's learning process that has a negative binomial distribution.

A negative binomial distribution is the probability distribution of the number of failures before the r^{th} success in a Bernoulli process with probability ρ of success on each trial. In other words, for t independent and identically distributed Bernoulli trials with success probability ρ, it gives the probability of s successes and $t-s$ failures, with success on the last trial. In our case, $NegBin(s^*, \rho) = NegBin(t_{GC}^*, \rho)$ distribution gives the probability of $t - t_{GC}^*$ failures and t_{GC}^* successes in t $Bernoulli(\rho)$ trials with success on the last trial, $t \geq t^*_{GC}$. The underlying random variable is the number of periods and its probability mass function is given by:

$$f(t) = \begin{pmatrix} (t - t^*_{GC}) + t^*_{GC} - 1 \\ t^*_{GC} - 1 \end{pmatrix} \rho^{t^*_{GC}} (1-\rho)^{t - t^*_{GC}} = \begin{pmatrix} t - 1 \\ t^*_{GC} - 1 \end{pmatrix} \rho^{t^*_{GC}} (1-\rho)^{t - t^*_{GC}}$$

where $\begin{pmatrix} t - 1 \\ t^*_{GC} - 1 \end{pmatrix} = \frac{\Gamma(t)}{(t^*_{GC}-1)!.\Gamma(t-t^*_{GC}+1)} = \frac{(t-1)!}{(t^*_{GC}-1)!(t-t^*_{GC})!}$. Function $f(t)$ gives the probability of t^*_{GC} successes in t periods with a success in t. In our context, the EA makes a proposal at even periods and $f(t)$ gives the probability of finishing the game at an even period $t \in t^*_{GC} + 2i$, $i = 0, 1, 2...$ if t^*_{GC} is even. If t^*_{GC} is odd, the game ends at $t \geq t^*_{GC} + 1$, $t \in t^*_{GC} + 1 + 2i$ with probability $f(t)$ because agreement never occurs at odd periods. Moreover, the probability that the game finishes on or before reaching any period $\widehat{t} > t^*_{GC}$ is given by $\sum_{i=t^*_{GC}}^{\widehat{t}} f(i)$. The number of successes in t periods with probability ρ of success on each trial is, on average, given by $t\rho$. Then, the t^*_{GC} successes take $\frac{t^*_{GC}}{\rho}$ number of periods on average.

If $t^*_{GC} = 0$, this means that it is not optimal to delay even one period, so the game ends immediately. ■

The following corollary shows that there exists a threshold level of learning probability such that if the probability is less than this level, the game ends immediately.

Corollary 1 *For any given model parameters β, δ, μ and λ_0, there exist an information processing probability ρ, ρ^*, such that whenever $\rho \leq \rho^*$, immediate agreement occurs.*

Proof. Assume that we are at time t. Then, $F(t; \beta, \delta, \mu, \lambda_0, \rho) = x^t_t(1-\delta^2(1-\rho)^2) - \delta^2\rho(\rho x^{t+2}_{t+2} + 2(1-\rho)x^{t+1}_{t+2})$. If we take derivative of $F(.)$ with respect to ρ, we get

$$\begin{aligned} \frac{dF}{d\rho} &= 2\delta^2(1-\rho)x^t_t - 2\delta^2\rho x^{t+2}_{t+2} - 2\delta^2(1-2\rho)x^{t+1}_{t+2} \\ &= 2\delta^2((1-\rho)x^t_t - \rho x^{t+2}_{t+2} - (1-2\rho)x^{t+1}_{t+2}) \end{aligned}$$

I argue that this is negative

$$\frac{dF}{d\rho} = 2\delta^2((1-\rho)x^t_t - \rho x^{t+2}_{t+2} - (1-2\rho)x^{t+1}_{t+2}) < 0 \qquad (2)$$

In order to show this, we make the following manipulations: since $x^t_t < x^{t+1}_{t+2} < x^{t+2}_{t+2}$,

$$\begin{aligned} \frac{dF}{d\rho} &= 2\delta^2((1-\rho)x^t_t - \rho x^{t+2}_{t+2} - (1-2\rho)x^{t+1}_{t+2}) \\ &< 2\delta^2((1-\rho)x^t_t - \rho x^{t+1}_{t+2} - (1-2\rho)x^{t+1}_{t+2}) \\ &= 2\delta^2((1-\rho)x^t_t - (1-\rho)x^{t+1}_{t+2}) \\ &= 2\delta^2((1-\rho)(x^t_t - x^{t+1}_{t+2}) < 0 \end{aligned}$$

Thus, $\frac{dF}{d\rho}$ is negative.

When $\rho = 0$, there will be an immediate agreement, $t^*_{GC} = 0$, since $F(s; \beta, \delta, \mu, \lambda_0, 0) > 0$. On the other hand, when $\rho = 1$, $F(s; \beta, \delta, \mu, \lambda_0, 1) = x^s_t - \delta^2 x^{s+2}_{t+2}$ may well be negative implying delay of the game, $t^*_{GC} \geq 0$. In addition, since $\frac{dF}{d\rho} < 0$, there must be a level of ρ, ρ^*, such that for all $\rho \leq \rho^*$, there would be an immediate agreement. Note that if $F(s; \beta, \delta, \mu, \lambda_0, 1) = x^s_t - \delta^2 x^{s+2}_{t+2} > 0$, then $\rho^* = 1$.

Another way to show this is as follows. By implicit function theorem, we can write

$$\frac{ds}{d\rho} = -\frac{\frac{dF}{d\rho}}{\frac{dF}{ds}} \text{ with } \frac{dF}{ds} \neq 0$$

by treating s as a continuous variable. Since $\frac{dF}{d\rho} < 0$, $sign(\frac{ds}{d\rho}) = sign(\frac{dF}{ds})$.

$$\frac{dF}{ds} = (1 - \delta^2(1-\rho)^2)\frac{dx^s_t}{ds} - \delta^2\rho^2\frac{dx^{s+2}_{t+2}}{ds} - 2\delta^2\rho(1-\rho)\frac{dx^{s+1}_{t+2}}{ds}$$

By using $0 < \frac{dx^{s+2}_{t+2}}{ds} < \frac{dx^{s+1}_{t+2}}{ds} < \frac{dx^s_t}{ds}$, it is easy to show that $\frac{dF}{ds} > 0$. Thus, $\frac{ds}{d\rho} > 0$. Then, the fact that s is a positive integer ($s(\rho)$ is actually a step function) and a similar argument above necessitates a threshold level of ρ. ■

In words, the opponent of the PNHA realizes that the continuation payoff increases with the learning probability. If the learning probability is too low, it is not worth to delay the game. Incentive for delaying increases as the learning probability gets closer to one (given the other model parameters). This obviously implies a threshold level for this probability.

On the other hand, independence of the state transitions from the state itself (the underlying random variable is i.i.d.) is crucial because ex ante, there is an integer t^*_{GC} referring to the number of successes (number of periods that the agent processes information correctly) but ex post, regardless of what happened in previous periods, the game looks the same from the perspective of the EA in terms of the incentives to delay the game.

The immediate agreement and delay results (like corollary 1) can be written in terms of other parameters of the model as well but since the focus here is on the parameter of information processing, we state only this result.

4.2 Perfect Bayesian Case

We now examine a case where the agent is a perfect Bayesian learner such that she processes available information perfectly. In other words, we impose $\alpha_1 = \alpha_2 = 1$. Given that the actual state of the world is the hyperbolic state, the PNHA updates her beliefs on the state of the world as the information flows during the course of the game.

As mentioned at the beginning, the exponential player is aware of the behavioral characteristics of the PNHA and plays accordingly. The PNHA is not aware of this own characteristic because we think that this type of behavioral characteristics of the PNHA naturally arises and this is how things are coded in her brain.[20]

Naive hyperbolic player updates her prior probability as she observes the outcome of each stage game. How long the game lasts will depend on the parameter values. As long as the game continues, the agent will face rejections. This causes belief updating that occurs as follows: The prior probabilities are λ_0 and $1-\lambda_0$ for the hyperbolic and mixed states, respectively. Let λ_t denote the belief at time t if the game is played t periods. A very simple Bayesian updating procedure indicates that $\lambda_t = \frac{\lambda_0}{\lambda_0+(1-\lambda_0)\mu^t}$. If we let $t \in \mathbb{R}$, then we get intuitive results $\frac{d\lambda_t}{dt} > 0$ and $\frac{d^2\lambda_t}{dt^2} < 0$. This means that the probability of hyperbolic state increases overtime at a decreasing rate. Then, the effective discount factor is again given by $\delta_s = \beta\delta + \mu\delta(1-\beta)(1-\lambda_s)$.

Given this preference profile and the information structure mentioned earlier, the following result is obtained that is similar to the main theorem in Akin (2007) with the exception of the difference in updating processes.

Proposition 1 In the game played between an exponential player and a partially naive player, in any stationary NBI solution, there exists a date t^* such that the players do not reach an agreement before t^* and at each time $t \geq t^*$, the players reach an agreement immediately at date t when the exponential agent offers; they reach an agreement at date $t+1$ when the partially naive agent offers.

Proof. This proof is identical with the proof of the theorem in Akin (2007) with the exception of modeling approaches to learning. In Akin (2007), the agent believes that using $\beta\delta$ discount factor in the future is a random variable distributed with some unknown parameter η, which measures the probability of the agent using $\beta\delta$ at any date t, and η is distributed with a beta distribution. In the current model, the beliefs are on the states of the world in which the agent is either in hyperbolic state or in mixed state. The updating structures differ due to these underlying models. In the proof, these two models only make the effective discount rates different but the logic and the steps are identical. ■

Corollary 2 $t^* \geq t^*_{GC} \geq 0$.

Proof. The function defined in Theorem 1, $F(s;\beta,\delta,\mu,\lambda_0,\rho)$, determines the critical value t^*_{GC} in generic case. The critical value in the perfect Bayesian case, t^*, is determined by the same function in which $\rho = 1$ since $\rho = \theta\alpha_1 + (1-\theta)\alpha_2$ and $\alpha_1 = \alpha_2 = 1$ in this case. By Corollary 1, we know that $\frac{ds}{d\rho} > 0$. This implies that $t^* > t^*_{GC}$ for continuous values of s. For discrete values of s, we have to write $t^* \geq t^*_{GC}$. By definition, t^* and t^*_{GC} may well be zero. Thus, $t^* \geq t^*_{GC} \geq 0$. ■

This corollary indicates that *the number of periods of delay* in perfect Bayesian case is at least as large as *the number of successful learning periods* in imperfect information processing case. This is intuitive when one

[20] The issue of more sophisticated behavioral players is beyond the scope of this paper and is left for future research. For a model of self deception and endogenous memory, see Benabou and Tirole (2002).

thinks about the trade off the EA faces. Incentive to delay the game increases with the learning probability. However, this does not mean that the delay in the former will always be more than the one in the latter. The reverse may well be the case.

4.3 Sudden Learning

In this section, we will assume that $\mu = 1$. This means that there are two mutually exclusive states of the world, exponential or hyperbolic. In other words, whenever there is a rejection and this information is processed (correctly) by the PNHA, it is revealed that the state of the world is the hyperbolic state (she becomes sophisticated). This occurs with a fixed probability at each $t \geq 0$ (sudden learning) and after being sophisticated, the agent stays sophisticated. The fixed probability can be defined as $\rho = \theta\alpha_1 + (1-\theta)\alpha_2$ based on the mentioned framework. The following proposition characterizes the relationship between the equilibrium structure of the bargaining game and the fixed probability ρ.[21]

Proposition 2 *Let an EA and a PNHA play the alternating-offers bargaining game with the specifications of this section. Then, in any stationary NBI solution, for any $\delta \in (0,1)$, $\beta \in (0,1)$ and $\widehat{\beta} \in (\beta, 1)$, there exists a learning probability ρ^* such that for all $\rho \leq \rho^*$, the EA offers $\underline{S}$ and agreement occurs immediately. For all $\rho > \rho^*$, the EA offers $\overline{S}$ in which the agreement is delayed with probability $1-\rho$. When the PNHA offers, either, with probability ρ, she offers $\delta\overline{S}$ to the EA and the game ends immediately or, with probability $1-\rho$, she offers $\widetilde{S}$ to the EA and the agreement is delayed.*

Proof. First remember the previously stated definitions: $\overline{S} = \frac{1-\beta\delta}{1-\beta\delta^2}$, $\underline{S} = 1 - \beta\delta\frac{1-\delta}{1-\widehat{\beta}\delta^2}$ and $\widetilde{S} = \frac{\delta-\widehat{\beta}\delta^2}{1-\widehat{\beta}\delta^2}$. Take any $\delta \in (0,1)$, $\beta \in (0,1)$ and $\widehat{\beta} \in (\beta, 1)$. Now define two value functions for the EA's problem as follows:

$$V(t+1) = \max_{\{A,\ R\}} \{\delta\overline{S}, \delta V(t+2)\},$$

$$V(t) = \max_{\{\underline{S},\overline{S}\}} \{\underline{S}, E(\overline{S})\},\ \textit{where}\ t \in \{0,2,4,...\}.$$

At each even period, the EA either offers $\overline{S}$ or $\underline{S}$ because the opponent is either sophisticated in which case the EA can get $\overline{S}$ or partially naive in which case the EA can get $\underline{S}$. He does not offer anything between $\overline{S}$ and $\underline{S}$ because a partially naive opponent rejects this offer anyway and a sophisticated opponent is ready to accept $\underline{S}$. If he offers $\underline{S}$, both types accept for sure. If he offers $\overline{S}$, this is accepted with probability ρ and rejected with probability $1-\rho$ and his continuation payoff is $V(t+1)$. Thus, the expected value of offering $\overline{S}$, $E(\overline{S})$, is $\rho\overline{S} + (1-\rho)\delta V(t+1)$.

[21] Mood states can take simpler forms such as the PNHA may stay in the same mood during the whole game, or information processing probabilities may be equal or one of the information processing probabilities may equal to zero. At each case, the fixed probability, ρ, can be defined in terms of model parameters $(\theta, \alpha_1, \alpha_2)$. If she stays in the same mood during the whole game, given the mood i, $\rho = \alpha_i$. If the information processing probabilities are equal, $\rho = \alpha_1 = \alpha_2$. If one of the information processing probabilities is equal to zero, then $\rho = (1-\theta)\alpha_2$ if $\alpha_1 = 0$, $\rho = \theta\alpha_1$ if $\alpha_2 = 0$.

At each odd period, the EA either accepts or rejects the opponent's offer. The opponent is either sophisticated in which she offers $\delta\overline{S}$ that is accepted or still partially naive in which she offers $\widetilde{S}$. By lemma 1 in Akin (2007), offers of the PNHA are never accepted. In other words, the game can end only at even periods in which the EA offers unless the PNHA becomes an SHA at an odd period. The immediate reward function at odd periods takes two values $\delta\overline{S}$ or $\widetilde{S}$. However, unless the offer is $\delta\overline{S}$, the EA rejects the offer with a continuation payoff $V(t+2)$. Thus, we can write the followings:

$$\begin{aligned} E(\overline{S}) &= \rho\overline{S} + (1-\rho)\delta V(t+1) \\ V(t+1) &= \rho\delta\overline{S} + (1-\rho)\delta V(t+2) \\ V(t) &= V(t+2) \end{aligned}$$

The value function takes the form of $V(t)$ for even periods and $V(t) = V(t+2)$ because of the static environment in which the opponent's type is exogenously determined with a fixed probability. Then,

$$V(t) = \max_{\{\underline{S},\overline{S}\}} \{\underline{S}, E(\overline{S})\}$$

$$V(t) = \max_{\{\underline{S},\overline{S}\}} \{\underline{S}, \rho\overline{S} + (1-\rho)\delta(\rho\delta\overline{S} + (1-\rho)\delta \max_{\{\underline{S},\overline{S}\}} \{\underline{S}, E(\overline{S})\})\}$$

Now, if $E(\overline{S}) > \underline{S}$, which means that the EA offers $\overline{S}$, then

$$E(\overline{S}) = \rho\overline{S} + (1-\rho)\delta(\rho\delta\overline{S} + (1-\rho)\delta E(\overline{S})) \tag{3}$$

If $E(\overline{S}) \leq \underline{S}$, which means that the EA offers $\underline{S}$, then

$$E(\overline{S}) = \rho\overline{S} + (1-\rho)\delta(\rho\delta\overline{S} + (1-\rho)\delta\underline{S}) \tag{4}$$

These two expressions should be consistent in the sense that $E(\overline{S}) > \underline{S}$ or $E(\overline{S}) \leq \underline{S}$ is satisfied where $E(\overline{S})$ is defined in (3) and (4), respectively. By using $E(\overline{S}) > \underline{S}$ and (3), we get:

$$H(\delta, \beta, \widehat{\beta}; \rho) = \frac{\rho\overline{S}(1 + (1-\rho)\delta^2)}{1 - \delta^2(1-\rho)^2} - \underline{S} > 0 \tag{5}$$

By using $E(\overline{S}) \leq \underline{S}$ and (4), we get:

$$H(\delta, \beta, \widehat{\beta}; \rho) \leq 0 \tag{6}$$

Note that $\frac{dH(\delta,\beta,\widehat{\beta};\rho)}{d\rho} > 0$, $H(\delta, \beta, \widehat{\beta}; 0) < 0$ and $H(\delta, \beta, \widehat{\beta}; 1) > 0$ since $\overline{S} > \underline{S}$. Thus, there must be a threshold value ρ^* satisfying $H(\delta, \beta, \widehat{\beta}; \rho^*) = 0$ such that for all $\rho \leq \rho^*$, (6) is satisfied. Otherwise, (5) is satisfied. Thus, if $\rho \leq \rho^*$, the EA offers $\underline{S}$ in which case the game ends immediately and if $\rho > \rho^*$, he offers $\overline{S}$ in which case the game is delayed with probability $1-\rho$. When the PNHA offers, she is still naive with probability $1-\rho$ in which case she offers $\widetilde{S}$ and the game is delayed. She is sophisticated with probability ρ in which case she offers $\delta\overline{S}$ and the game ends immediately. ■

Corollary 3 *If $\rho > \rho^*$, then the game is delayed $\widetilde{t}$ periods with probability $\rho(1-\rho)^{\widetilde{t}}$.*

Proof. Suppose $\rho > \rho^*$. By Proposition 2, we know that as long as the PNHA does not become an SHA, the game is delayed. Then, a delay of $\widetilde{t}$ periods has a probability of $(1-\rho)^{\widetilde{t}}$. Moreover, the game ends at period $\widetilde{t}+1$ with probability ρ. Thus, the game is delayed $\widetilde{t}$ periods and ends at $\widetilde{t}+1$ with probability $\rho(1-\rho)^{\widetilde{t}}$. ■

The proposition characterizes a threshold level for the learning probability above which it is optimal to treat the opponent as a sophisticated agent for the EA. The corollary specifies probabilistically when the bargaining game ends.

4.3.1 Comparative Statics

We now find how the values of ρ^* change with different parameters of the model by applying implicit function theorem on $H(\delta, \beta, \widehat{\beta}; \rho^*)$. We can explore the relationship between ρ^* and $\widehat{\beta}$ by examining the following partial derivative:

$$\frac{\partial \rho^*}{\partial \widehat{\beta}} = -\frac{\frac{dH(.)}{d(\widehat{\beta})}}{\frac{dH(.)}{d(\rho^*)}} < 0$$

Note that, $\frac{dH(.)}{d(\rho^*)} > 0$ and $\frac{dH(.)}{d(\widehat{\beta})} > 0$ for all parameter values. This means that the EA may offer $\overline{S}$ to a more naive opponent whom he does not offer $\overline{S}$ originally. In other words, against a more naive agent (a higher $\widehat{\beta}$), the EA continue to offer $\overline{S}$ even the agent's learning probability is lower. The intuition behind this is that against a more naive opponent, the payoff that can be guaranteed by the EA is lower since the minimum share that the more naive agent can accept is higher. Thus, against an agent with a higher $\widehat{\beta}$, the EA is more willing to take a risk and offer $\overline{S}$.

In order to examine the relationship between ρ^* and β, we need the following partial derivative:

$$\frac{\partial \rho^*}{\partial \beta} = -\frac{\frac{dH(.)}{d(\beta)}}{\frac{dH(.)}{d(\rho^*)}} < 0$$

$\frac{dH(.)}{d(\beta)}$ cannot be assigned any sign analytically but based on some simulations, we found that $\frac{dH(.)}{d(\beta)} > 0$, and since $\frac{dH(.)}{d(\rho^*)} > 0$, we get $\frac{\partial \rho^*}{\partial \beta} < 0$. This means that as the self-control problem of the agent lessens (a higher β), the EA is more likely to offer $\overline{S}$. As β increases, both $E(\overline{S})$ and $\underline{S}$ decrease but $\underline{S}$ decreases more than $\overline{S}$. This leads to an increase in tendency to offer $\overline{S}$ and this means a lower ρ^*.

5 Discussion and Conclusion

Identifying behavioral biases of people, determining implications of them and then, if possible, suggesting operational policies are extremely relevant and crucial issues since people frequently exhibit bounded, rather than perfect, rationality. In this paper, we investigate the behavior of agents in a strategic environment who potentially have two well known biases. Since this is a theoretical but not an experimental paper, we do not identify potential biases. We basically try to extract implications of some imposed biases that are observed frequently. This helps us understand the behavior of players who try to take advantage of other players having these biases.

We believe that having a preference for immediate gratification emerges as a very important factor and it gives surprising and seemingly irrational outcomes we face both in individual decision making and in strategic environments. Savings and investment behavior, project completion, intertemporal individual decision making problems, price determination, several advertorial strategies employed by firms and bargaining situations are some examples.

Moreover, since the very nature of acting rationally necessitates being aware of what the agent desires and how she optimizes under both feasibility and behavioral constraints, learning in the sense of increasing awareness is also a crucial part of rationality. This is the main reason why we choose imperfect information processing along with the present bias among the many biases presented both in economics and psychology literature. In addition, they together elucidate the intuition behind the persistency of implied anomalous behavior. This point indicates the fact that one bias may wash out the effect of the other or it may exacerbate the other's effect[22]. In our context, being naive about own preferences is a bias that makes the agent better off in a bargaining setting but learning own preferences overtime by experience alleviates the positive effect of this bias. On the other hand, imperfect information processing (due to confirmation bias) is another bias and in the existence of both of these biases, the latter prevents the lessening of the effect of the former. In this sense, one bias does not exacerbate the effect of the other but it provides permanence of the other's effect. Our mood state approach turns out to be a formal explanation of this persistency, e.g., imperfect information processing causes the present bias to be persistent that actually makes the agent better off. Although restricting the bias space into these two biases may seem as a limitation, this result can be stated for any bias combinations having similar implications.

While we find that being naive about own (present biased) preferences is advantageous, most of the industrial organization literature finds that it hurts the decision maker (e.g., Dellavigna and Malmendier, 2004; Sarafidis, 2005). In these environments, we observe persistency of different biases (Loewenstein, 1996; Van Boven, Dunning and Loewenstein, 1999) and people do not / cannot correct these biases although having

[22] Besharov (2004), by using examples of time inconsistency, regret, and overconfidence, examines how biases may offset each other's effects and the implications of correcting biases.

them is costly. In this case, permanence of them is less likely. In this paper's framework, however, people actually get better off by failing to appreciate the bias they have, hence they do not have any incentive to overcome this bias in the sense of learning and becoming more aware of it[23]. Hence, it is more likely to see a permanence of the bias. Moreover, even if they become more aware overtime, they have an incentive to imitate to be naive. In this sense, examining alternating-offers bargaining game as the medium of interaction is more interesting in terms of its implications[24], but it is possible to extend this work by considering different types of games.

In this paper, we model information processing in a specific way by assuming that the number of learning period is a random variable that has a negative binomial distribution. This is actually a simple way to model the behavior we focus on and we conjecture that even if it is assumed different stochastic learning structures for learning, the results do not change significantly. However, the learning focused on here is better to be understood as case by case learning (This can be called as meta-learning). It is difficult to talk about learning in a global sense meaning that agents become fully aware of a bias they have regardless of the decision making problem.

As Babcock et.al. (1995) argues, people may recognize self-serving biases of other people but they tend to overlook the ones they may have. Rabin and Schrag (1999) raised a question of how economic implications might depend on people's awareness of others' confirmatory bias and they continue:

> "One possibility is that people might exploit the bias of others. A principal may, for instance, design an incentive contract for an agent that yields the agent lower wages on average than the agent anticipates, because the agent will be overconfident about her judgments in ways that may lead her to exaggerate her yield from a contract."

We assume that the principal is aware of the existence of the biases the agent has and acts accordingly[25]. Given this assumption, our mood state approach naturally gives rise to some policy issues from the perspective of the principal, namely, mood induction. A numerous strategies used by producers (suppliers of goods and services in general) can be seen as mood induction. Presenting pre-funding options, bundling of several services under the same contract, pointing out tempting characteristics of products, supermarkets having stands for foods to try and slow music in the background, car dealers offering test drives, real estate agencies showing sample houses, free trial periods of cable TV's and cell phones are some examples of strategic

[23] In especially intertemporal games, being naive refers to the concept that we call as "believing is as if actually being" under, of course, some informational assumptions.

[24] Other reasons of choosing alternating-offers bargaining game were mentioned in the introduction (it is a sequential, infinitely played, two person, and dynamic game that has a wide range of real life applications and a well established theory).

[25] This example overlaps almost perfectly with the example of this paper which is the bargaining game. The principal's design is (wage) bargaining game and the naive agent in advance anticipates a higher wage but it turns out to be less than what she anticipated (as a result of learning).

behavior of suppliers to exploit the bias of consumers. These are examples of principal trying to activate the agent's hot system in the sense of Metcalfe and Michel (1999). On the other hand, in our context, the principal try to induce moods in which processing information is more likely (activating cool system) because information processing makes the principal better off.

In conclusion, this paper is an initial step to investigate the behavioral characterization of agents having self-serving biases. However, extending this into more general settings still remains an issue that needs further research. Possible extensions are to relax the informational assumptions, to allow for correlated cognitive and mood states, to generalize the learning dynamics (especially, to allow for two or more conflicting signals in a different game setting that result in ambiguous dynamics of beliefs as in Rabin and Schrag, 1999), to relax the information processing structure and to consider different types of strategic environments.

Acknowledgement

I thank Burak Erdeniz, who is the research assistant in this project, for his efforts. For helpful comments, I also thank Ahu Genis Gruber, Derek Laing, Emin Karagozoglu, Isa Emin Hafalir and Thomas Sjostrom. All remaining errors are mine.

REFERENCES

Akin, Z. (2004). *The Role of Time-Inconsistent Preferences in Intertemporal Investment Decisions and Bargaining.* Mimeo Pennsylvania State University.

Akin, Z. (2007). Time Inconsistency and Learning in Bargaining Games. *International Journal of Game Theory*, 36:2,275-299.

Antonides, G. (1996). *Psychology in Economics and Business : An Introduction to Economic Psychology.* Kluwer Academic Publishers, Dordrecht, (Chapter 7).

Babcock, L. & Loewenstein, G. (1997). Explaining bargaining impasse: the role of self-serving biases. *Journal of Economic Perspectives*, 11, 109-126.

Babcock, L., Loewenstein, G., Issacharoff, S. & Camerer, C. (1995). Biased Judgments of Fairness in Bargaining. *American Economic Review*, 85-5, 1337-1343.

Bechara, A., Damasio, A.R., Damasio, H. & Anderson, S.W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, 50, 7-12.

Benabou, R. & Tirole, J. (2002). Self-Confidence and Personal Motivation. *Quarterly Journal of Economics*, 117:3, 871-915.

Bernheim, B. D. & Rangel, A. (2004). Addiction and Cue-Triggered Decision Processes. *American Economic Review*, 94:5, 1558-1590.

Besharov, G. (2004). Second-Best Considerations in Correcting Cognitive Biases. *Southern Economic Journal*, 71:1, 12-20.

Brocas, I. & Carrillo J. D. (2004). *Biases in Perceptions, Beliefs and Behavior*, Discussion paper, University of Southern California.

Chade, H., Prokopovych, P. & Smith, L., (2008). Repeated games with present-biased preferences. *Journal of Economic Theory*, 139, 157-175.

Compte, O. & Postlewaite, A. (2008). *Repeated Relationships with Limits on Information Processing.* PIER Working Paper, 08-026.

Damasio, A.R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain.* New York: Grosset/Putnam.

Dellavigna, S. & Malmendier, U. (2004). Contract Design and Self-Control: Theory and Evidence. *Quarterly Journal of Economics*, 119:2, 353-402.

Dellavigna, S. & Malmendier, U. (2006). Paying Not To Go To The Gym. *American Economic Review*, 96, 694-719.

Ellis, H. C. & Ashbrook, P. W. (1988). Resource allocation model of the effects of depressed mood states on memory. K. Fiedler & J. P. Forgas (Eds.), *Affect, Cognition, and Social Behavior*, 25–43, Toronto: Hogrefe.

Fudenberg, D. & Levine, D. K. (2006). *A Dual Self Model of Impulse Control.* Harvard Institute of Economic Research Discussion Paper, No. 2112.

Isen, A. M. (2002). A role for neuropsychology in understanding the facilitative influence of positive affect on social behavior and cognitive processes. Snyder, C. R., & Lopez, S. J. (Eds.), *Handbook of Positive Psychology*, 528-540, NY:Oxford.

Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics. *American Economic Review*, 93:5, 1449-1475.

Koszegi, B. (2000). *Ego Utility, Overconfidence and Task Choice.* Unpublished Paper.

Laibson, D. (1997). Golden eggs and Hyperbolic Discounting. *Quarterly Journal of Economics*, 62:2, 443-478.

Lipman, L. B. (1995). Information Processing and Bounded Rationality: A Survey. *The Canadian Journal of Economics / Revue canadienne d'Economique*, 28:1, 42-67.

Loewenstein, G. (1996). Out of Control: Visceral Influences on Behavior. *Organizational Behavior and Human Decision Processes*, 65, 272-92.

Loewenstein, G., O'Donoghue, T. & Rabin, M. (2003). Projection Bias in Predicting Future Utility. *Quarterly Journal of Economics*, 1029 - 1248.

Maia, T. V. & McClelland, J. L. (2004). A reexamination of the evidence for the somatic marker hypothesis: What participants really know in the Iowa gambling task. *Proceedings of National Academy of Sciences*, 101:45, 16075–16080.

McClure, S. M., Laibson, D., Loewenstein, G. & Cohen, J. D. (2004). Separate Neural Systems Value Immediate and Delayed Monetary Rewards. *Science*, 306:5695, 503 - 507.

McGraw, A. P., Mellers, B. A. & Ritov, I. (2004). The Affective Costs of Overconfidence. *Journal of Behavioral*

Decision Making, 17: 281–295.

Metcalfe, J. & Michel, W. (1999). A hot/cool system analysis of delay of gratification: Dynamics of willpower. *Psychological Review*, 106:1, 3-19.

Myers, D. G., *Social Psychology,* 1996, McGraw-Hill, New York.

Neale, M. A. & Bazerman, M. H. (1985). The Effects of Framing and Negotiator Overconfidence on Bargaining Behaviors and Outcomes. *The Academy of Management Journal*, 28:1, 34-49.

O'Donoghue, T. & Rabin M. (1999a). Doing It Now or Later. *American Economic Review*, 89:1, 103-124.

O'Donoghue, T. & Rabin M. (2001). Choice and Procrastination. *Quarterly Journal of Economics*, 116:1, 121-160.

Pinto, L.S. & Sobel, J. (2005). A Model of Positive Self-Image in Subjective Assessments. *American Economic Review*, 95:5, 1386-1402.

Rabin, M. & Schrag, J. L. (1999). First Impressions Matter: A Model of Confirmatory Bias. *Quarterly Journal of Economics*, 114:1, 37-82.

Raghunathan, R. & Pham, M. T. (1999). All Negative Moods Are Not Equal: Motivational Influences of Anxiety and Sadness on Decision Making. *Organizational Behavior and Human Decision Processes*, 79:1, 56–77.

Rubinstein, A. (1982). Perfect Equilibrium in a Bargaining Model. *Econometrica*, 50, 1, 97-110.

Sanbonmatsuand, D. M. & Kardes, F. R. (1988). The Effects of Physiological Arousal on Information Processing and Persuasion. *Journal of Consumer Research*, 15:3, 379-385.

Sarafidis, Y. (2005). *Inter-temporal Price Discrimination with Time-inconsistent Consumers.* Mimeo CRA International.

Schwarz, N., Bless, H., & Bohner, G. (1991). Mood and persuasion : affective states influence the processing of persuasive communications. *Advances in experimental social psychology*, 24, 161-199.

Schwarz, N. & Clore, G.L. (1983). Mood, misattribution, and judgments of well-being: informative and directive functions of affective states. *Journal of Personality and Social Psychology*, 45:3, 513-523.

Schwarz, N. & Clore, G.L. (2003). Mood as information: 20 years later. *Psychological Inquiry*, 14, 294-301.

Schwarz, N. (2000). Emotion, cognition, and decision making. *Cognition and Emotion*, 14:4, 433-440.

Seligman, E., (1990). *Learned Optimism: How to Change Your Mind and Your Life.* New York: Simon and Schuster.

Stanovich, K. E. & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, 23, 645-665.

Strotz R. H. (1956). Myopia and inconsistency in Dynamic Utility Maximization. *Review of Economic Studies*, 23, 165-180.

Van Boven, L., Dunning, D. & Loewenstein, G. (1999). *Trading places: egocentric empathy gaps between owners and buyers.* Unpublished manuscript, Cornell University.

Wason, P.C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, 12, 129-140.

Zabojnik, J., (2004). A Model of Rational Bias in Self-Assessments. *Economic Theory*, 23:2, 259–282.