Limited Foresight Equilibrium

Jeevant Rampal*

August 17, 2017

Abstract

This paper defines the Limited Foresight Equilibrium (LFE). Foresight is defined as the number of subsequent stages of a sequential game that a player can observe from a given move. In the context of a finite sequential game with perfect information, we model a scenario where players can possess various levels of limited foresight and each player is uncertain about her opponents' foresight-levels. The LFE provides an equilibrium assessment for this model. We show the existence of LFE. In LFE, limited-foresight players' perception of the game changes as they move through the stages of the game; their strategies evolve and they update their beliefs about the opponents' foresights within the play of the game. If a player has greater foresight, then her LFE beliefs about the opponents' foresights are more accurate. If a limited-foresight player finds herself at an "unexpected" position, she discovers that she is playing against some higher foresight opponent. Players' LFE strategies take reputations about their foresight into account. In applications, LFE is shown to rationalize experimental findings on the Bargaining game and the Centipede game. The LFE's novel predictions are corroborated by data from a modified Race game.

Keywords: Limited Foresight, Sequential Equilibrium, Bargaining, Centipede Game.

JEL Codes: C72, C78, D83, D91.

*Indian Institute of Management Ahmedabad. Email: jeevantr@iima.ac.in. Part of this paper was written at the Ohio State University. I thank James Peck for his encouragement, invaluable comments and guidance in both conceptualizing and writing this paper. I thank Larbi Alaoui, Yaron Azrieli, A. Banerji, Paul Healy, Philippe Jehiel, Anirban Kar, Shaowei Ke and Antonio Penta for their insightful comments and discussions. Last, I thank the numerous seminar participants who have given their helpful comments.

1 Introduction

Consider the setting of a finite set of players playing a sequential move game where each player knows all prior actions played at each of her moves. That is, consider the setting of a finite sequential game with perfect information. To solve for her optimal action at each move, each player should think ahead to the last stage and reason backwards using backward induction. That is, each player's optimal strategy should be calculated using the payoff possibilities in the game and her belief about the actions of her opponents in the future stages of the game. However, there are several experimental studies showing that even when this backward induction can reveal a weakly dominant strategy for the player, a high proportion of players are unable to perform such backward induction (cf. "Race Game" findings of Rampal (2017), Mantovani (2014), Levitt, List and Sadoff (2011)).[1] It is hard to explain this failure to play a dominant strategy using Dynamic Level-k models[2] (Ho and Su (2013), Kawagoe and Takizawa (2012)). These models posit that players are utility maximizers but their chosen strategies are determined by their subjective beliefs about their opponent's cognitive-level, which determines the opponent's strategy. However, such subjective beliefs don't help in explaining the failure to play weakly dominant strategies, precisely because a weakly dominant strategy is a *strict* best response *irrespective* of the players' subjective beliefs as long as players believe that their opponent(s) can play the perfect strategy with a strictly positive probability. Further, as Johnson et al (2002) show, a sizable proportion of players ignore payoff-relevant information about the future stages of a game even when it is available upon browsing on their decision screen. Therefore, the limited ability to think/look ahead in a multi-stage game appears to be a specific form of bounded rationality which generates corroborative patterns of behavioral data, for example,

[1]For example, in Rampal (2017), in one of the treatments, a highly significant 40 percent (approximately) of the subjects fail to play the dominant strategy in a "winner-take-all" sequential game against a computer which plays perfectly.

[2]Unless one resorts to a very specific strategy for the Level-0 player which involves failing to play the dominant strategy at the first few stages of a sequential game but playing perfectly towards the last few stages.

the tendency to make lower proportion of dominated choices as one gets closer to the end of a dynamic-game (cf. Rampal (2017), Mantovani (2014), Levitt, List and Sadoff (2011)).

We model this "limited ability to think/look ahead in a multi-stage game," or what we call *limited foresight*, as one of the two main components of our theory. Foresight is defined to be the number of subsequent stages that a player can look ahead in a multi-stage sequential game.[3] If the foresight of a player does not extend to the last stage of the game from each of her moves, she is said to have limited foresight.[4]

The second, novel, feature of this paper is that we also model the scenario where players are uncertain about their opponents' foresights. Several experimental studies (cf. Rampal (2017), Levitt, List and Sadoff (2011), and Palacios-Huerta and Volij (2009), among others) have combined players with different degrees of expertise in sequential games with perfect information and found that the information about the opponents' expertise has a significant impact on behavior.[5] We model the scenario where this degree of expertise is captured by the level of foresight. That is, we model the scenario where players are playing a "seemingly" perfect information game, which means that the players can observe all prior actions every time they move, but the players are uncertain about the levels of foresight of their opponents. That is, players appear to be playing a game with perfect information, but they are actually playing a game with imperfect information with uncertainty about the opponents' foresights. As a result, in our model, a limited foresight player's "optimal" choice depends on both his foresight and his belief about the opponents' foresights.[6]

[3]Other, independent, studies that have modeled limited foresight similar to this paper are Ke (2017), Mantovani (2014), and Roomets (2010).

[4]We do not model or investigate *why* players have limited foresight. For example, we don't answer the following questions. Does limited foresight occur because calculations are harder with more stages in the game? Or does it occur because players think that they don't need to consider future stages? Instead we focus on modeling only the implication of these possible primitives, i.e., limited foresight, and simultaneously we model uncertainty and belief updating about opponents' foresights. For an epistemic discussion of backward induction see Aumann (1995), Battigalli (1997), Ben-Porath (1997), Binmore (1996), and Brandenburger and Friedenberg (2014). Bonanno (2001) studies backward induction in terms of temporal logic.

[5]While Rampal (2017) induces different degrees of expertise by varying the degree of experience of the players in the game tested there, Levitt, List and Sadoff (2011) and Palacios-Huerta and Volij (2009) do so by mixing expert chess players with student subjects.

[6]Optimal choices also depend on his beliefs about the opponents' beliefs about his foresight, and so on, but we make the assumption that the prior distribution over foresight-levels across players is common

In our model, in addition to being uncertain about the opponents' foresights, players also update their beliefs about the opponents' foresights as they observe more moves of their opponents. That is, beliefs about the opponents' foresights evolve within the play of a game. The Limited Foresight Equilibrium (henceforth LFE) that we define and apply in this paper formalizes the meaning of "optimal" choices given "consistent" beliefs in this framework of limited foresight and uncertainty about the opponents' foresights.

The summary of the model and LFE is as follows. We start with an arbitrary game with "seemingly" perfect information. For example, consider Ann and Bob playing a Sequential Bargaining game (Rubinstein (1982) and Ståhl (1972)) where all prior choices are displayed. We map this game with perfect information to a game with imperfect information, called an *Interaction Game*. In our example, the Interaction Game is a scenario where there are multiple possible types of Ann and multiple possible types of Bob. Each type of Ann (respectively Bob) is uncertain about which type of Bob (Ann) she (he) is bargaining with. A type denotes a particular level of foresight.

Mapping the game with perfect information to its corresponding Interaction Game models the uncertainty about the opponent's type/foresight-level. However, this doesn't model limited foresight. In particular, to solve for the optimal actions of a limited-foresight type of Ann, we cannot use the whole Interaction Game because the limited-foresight type, by definition, cannot observe all the stages of the Interaction Game. Therefore we must consider curtailed versions of the Interaction Game to model both limited foresight and uncertainty about the opponent's foresight simultaneously. We call such curtailed versions of the Interaction Game as *Curtailed Games*.

The LFE is solved and defined recursively. We start with the shortest possible curtailed version of the Interaction Game: the 1-staged Curtailed Game(1), named $CG(1)$, in which

knowledge, which lets us end this line of thinking at the second-order beliefs. The last section of this paper weakens the common prior assumption to model subjective prior beliefs. Even in the subjective prior case, higher-than-second-order beliefs are not considered. The properties and evolution of these first-order and second-order beliefs within the play of a sequential game make for a "rich" analysis. Extension to models which deal with limited foresight and higher order beliefs is left for future research.

we curtail the Interaction Game after the first stage actions. If Ann starts the bargaining game, then $CG(1)$ is the game observed by the 0-foresight type of Ann, or 0_{Ann} at stage-1. This is because 0_{Ann} can observe 0 plus the stage number of her current move from each of her moves (in this case the stage number is 1). $CG(1)$ models both: (i) the fact that 0_{Ann} has a foresight-level of 0; and (ii) that 0_{Ann} is uncertain about Bob's type. We solve for the Sequential Equilibrium (henceforth SE) (Kreps and Wilson (1982b)) of $CG(1)$ to obtain the first stage LFE action and beliefs of 0_{Ann}.[7] Next, taking the stage-1 LFE action of 0_{Ann} as given as Nature's moves in $CG(2)$ (Curtailed Game(2)), we solve for the SE of $CG(2)$, thereby solving for the LFE actions and beliefs of 1_{Ann} at her stage-1 information set, and 0_{Bob} at his stage-2 information set; as they both observe $CG(2)$ at those information sets. We proceed recursively to obtain the LFE actions and beliefs for all the information sets of the Interaction Game. Therefore, in LFE, all foresight-types do the best they can within the bound of their foresight, given their belief about the probability distribution on opponents' types. The LFE accounts for the fact that the total foresight (defined as the total number of stages of the Interaction Game observed) and beliefs of a limited-foresight player evolve as she plays the Interaction Game.

Defining LFE as above provides us with the following properties. First, the LFE exists and it is upper hemi-continuous with respect to the prior distribution and payoff profile. Second, higher foresight types correctly anticipate lower types' moves. This property is approximately mirrored in a finding from Reynolds' (1992): testing recognition of opponent's expertise among chess players, he found that "Higher rated players consistently made lower

[7] To make $CG(1)$ a well defined game we need to define payoff profiles for its terminal histories/nodes. In particular, if we are curtailing an Interaction Game with more than two stages after the first stage actions, then how do we construct payoffs after the first stage actions in $CG(1)$? In this paper, in each Curtailed Game, after each last-stage action of that Curtailed Game, we construct a payoff profile such that the payoff to each player-type is equal to the $(min + max)/2$ of the payoffs possible for that player-type from that action in the Interaction Game. For example, if 0_{Ann} asks for the whole first stage pie (worth a 100 units) in the first stage, then curtailing the bargaining Interaction Game after she demands 100 in the first stage implies a payoff of $(100+0)/2$ in $CG(1)$. This is because the minimum Ann can get after making any offer is 0 (when there is no agreement among Ann and Bob at any stage), and the maximum is 100. We explain this "curtailment rule" in more detail in the model section. Mantovani (2014) also uses this same "curtailment rule" to map payoffs from a sequential game to payoffs for its curtailed versions. Ke (2017) formulates the axiomatic foundations of different possible "curtailment rules."

estimation errors" (of other chess players' ELO ratings). Third, when high-foresight types are estimating which lower-foresight opponent-type they are playing against, their belief becomes more accurate as they observe more moves of the opponent. This property is also mirrored in Rampal (2017) and Reynolds (1992); both studies found that the estimation error (about the opponent's expertise) decreased as the number of prior moves revealed increased. Fourth, due to the the recursive definition of LFE, if a player-type observes actions that were not part of the LFE strategies of opponents'-types with lower foresight than himself, then he discovers that he is playing against some opponent-type with higher foresight than himself. The attempt by limited-foresight types to recognize opponent-type and adjust behavior foreshadows the fifth feature of our model: reputation effects. High-foresight types have to decide on what's optimal: pretending to be a low type or revealing their high-foresight type. Importantly, in LFE, the updating of beliefs by different foresight-types happens *within a play of the sequential game.*

Two applications of LFE are described in this paper. In both applications we do not use game-specific types who follow specific strategies tailor made for the game; we only work with the tools of the LFE model: Limited foresight and uncertainty about the opponent's foresight. In the first application, LFE is shown to have the ability to explain several qualitative findings on the Sequential Bargaining game. We show that the LFE can simultaneously explain delays, near equal splits, disadvantageous counterproposals and subgame inconsistency. In particular, according to the LFE, disadvantageous counterproposals (Ochs and Roth (1989)) can be caused in a three period bargaining game because the second mover, when of a specific limited-foresight type, can fail to take into account that he has no bargaining power in the third period and that the pie shrinks after he rejects the first period offer. Thus, he rejects the first-period offer he receives, but when he has to make a counterproposal, he faces a shrunk second-period pie and he realizes his lack of last-period bargaining power, which make a disadvantageous counterproposal sequentially rational. In the second application, we show that the LFE implies passing by all types of both players (including the type with no foresight limitation) until the last few stages in a Centipede game with more than 4 stages, for arbitrary probability on limited foresight. This result follows from a reputation argument similar to (McKelvey and Palfrey (1992)) and Kreps, Milgrom, Roberts and Wilson (1982).

1.1 Experimental Findings Regarding the Limited Foresight Equilibrium: A Summary

For an experimental evaluation of the LFE model, we direct the reader to Rampal (2017) which uses "race games" to test the novel predictions of LFE. We summarize the findings of Rampal (2017) here. In the first experiment in Rampal (2017), the "race game" used there is as follows. Imagine a box containing 9 items. Players move alternately, removing 1, 2, or 3 items from the box at each move. The player who removes the last item loses, and his opponent wins. There is a second mover advantage in this game.[8] The prize from winning as the first-mover (second-mover) is 500 (200, respectively) experimental currency. Before the game begins, to decide the first and second movers, both players in a pair simultaneously choose between "first-mover" and "second-mover." Either of the two players' *choice* of starting position is selected and implemented with 50 percent chance each. Note that the prize from winning as the selected first-mover is greater than the prize from winning as the selected second-mover, but winning as the selected first-mover is possible *only if* the selected second-mover makes a mistake, that is, only if the selected second mover chooses a dominated action. Rampal (2017) found that if an experienced player, who had clearly displayed expertise in implementing the second-mover advantage in previous rounds, was told that his opponent was inexperienced then he was much more likely to choose "first-mover" as opposed to when the experienced player was told that his opponent was also experienced.

In the second (different) experiment in Rampal (2017) there are 13 items in the box which makes the game longer and the backward induction, required to solve for the SPNE strategy, harder. In this second experiment, experienced players were not told if their opponent was

[8] The second mover should remove ($4-$ opponent's previous choice) at each move to win.

experienced or inexperienced. Instead, an experienced player only observed how his opponent played a similar race game (with 13 items in the box) against a *perfectly playing computer*,[9] before the opponent and he played the race game (with 13 items in the box) against each other. The key finding from this experiment was that the experienced player was significantly more likely to choose "first mover" when he observed that his opponent had failed to play the dominant strategy against the computer, compared to the case where his opponent displayed perfect play against the computer. That is, Rampal (2017) found that both (i) exogenous information, and (ii) endogenous inference about the opponent's inexperience increase the probability with which experienced players abandon the "sure-win" strategy (of being the second-mover) and try for a higher payoff attainable only by winning from a losing position, i.e., a position from which one wins only if the opponent makes a mistake. A maximum likelihood analysis shows that compared to the AQRE model (McKelvey and Palfrey (1998)), the LFE model has a significantly higher likelihood with respect to the data in both experiments. The maximum likelihood analysis shows that compared to the Dynamic Level-k models (Ho and Su (2013) and Kawagoe and Takizawa (2012)), the LFE model generates a similar likelihood in the first experiment. In the second experiment, the number of items in the box is 13, which makes the sequential structure of the game more salient. This implies that the backward induction, required to understand and implement the selected second mover's dominant strategy, is harder. Therefore, the percentage of observations where the selected second mover fails to play the dominant strategy is higher compared to the first experiment (13.7% in the first experiment and approximately 57% in the second experiment). As these observations get 0 probability according to Dynamic Level-k models (unless an error structure is added to carry out the MLE), the LFE model

[9] In the race game between human and computer, the decision to go first or second was taken unilaterally by the human. The human subject was guaranteed a win against the computer if he/she followed the following strategy: Choose "second mover," and as the second mover, choose (4−computer's previous choice) at each move. Any deviation from this strategy meant a loss against the perfectly playing computer. The payoff from winning the race game versus the computer was an additional 500 experimental currency over and above the earnings from the race game versus human opponent. Thus, losing to the computer to signal low expertise meant losing 500 experimental currency. Thus, it was a dominated strategy.

has a significantly higher likelihood compared to the Dynamic Level-k models in the second experiment.

Therefore, three key aspects of the LFE model are established by Rampal (2017): first, the salience of the beliefs about the opponent's ability to do backward induction; second, that experienced players update their beliefs about the opponent's expertise based on the opponent's past moves in the game; third, that the LFE model has a significantly higher likelihood, with respect to this data in the second experiment, when compared to the AQRE model and the Dynamic Level-k models.

1.2 Related Literature

Most of the related Level-k literature deals with simultaneous move games (Stahl and Wilson (1995); Nagel (1995); Stahl (1996); Ho et al. (1998); Costa-Gomes et al. (2001); Camerer et al. (2004); Costa-Gomes and Crawford (2006); Crawford and Iriberri (2007a, b), among others), but it does capture the uncertainty about the opponent's expertise. The paper from this literature that is closest in spirit to LFE is Alaoui and Penta (2016) which endogenizes the choice of level in a Level-k framework. In their model, the choice of the level of a player is a function of his maximum possible level, which is endogenously determined by his incentives and costs of thinking about the game at hand, his belief about the opponent's maximum possible level and his belief about his opponent's belief regarding himself (i.e., his second-order beliefs). They disentangle the effect of these factors using a novel experimental design. While we don't model the analogous question of how a player's maximum possible foresight is determined, the LFE model does have the feature that a particular foresight type considers his first order beliefs (beliefs about opponents' foresight) and second order beliefs (his belief about the opponents' beliefs about his foresight) in choosing his optimal strategy. Further as we are dealing with sequential games, we also study how these beliefs evolve *across the stages of a single play of the game.* Experimental studies of the relation between the opponent's cognitive level and a player's choices in simultaneous Level-k settings include

Agranov et al (2012), Gill and Prowse (2014), and Slonim (2005), among others.

The closest works to ours in the area of sequential games are the working papers of Ke (2017), Mantovani (2014), and Roomets (2010). These papers, in independent projects, model limited foresight in a similar fashion to ours. Ke (2017) describes the axiomatic foundations of curtailment rules to model one player, multi-stage, decision problems as observed by a player with limited foresight. Mantovani (2014) endogenizes the choice of foresight. He also demonstrates the existence of limited foresight using an experiment on a different race game. The Valuation Equilibrium by Jehiel and Samet (2007) models how cognitively constrained players may group nodes of a sequential move game into exogenously given similarity classes, where each similarity class has a given valuation. Although related, Valuation Equilibrium does not deal with limited foresight specifically. Jehiel (1995) defines the Limited Forecast Equilibrium, where each player, at each of his moves, chooses his strategy to maximize his average payoff within his foresight horizon, given his forecast about the upcoming moves within that horizon. The forecasts are constrained to be consistent with the equilibrium strategies. Jehiel (1998a) provides a learning justification for these forecasts. Jehiel (1998b) and Jehiel (2001) extend the Limited Forecast Equilibrium to repeated games. The key difference among these papers and our work is that the second feature of our model, the uncertainty about the opponents' foresight-types, is absent from all the papers mentioned above. Consequently, updating belief about one's opponents' types within the game, strategic adjustments after updating these beliefs and reputation effects do not feature in these papers.

Ho and Su (2013) and Kawagoe and Takizawa (2012) have adapted the Level-k model for sequential move games. They allow for updating about the opponent's level/expertise *across repetitions of play* of the same game as opposed to *within* the play of a game as in here. The AQRE model of McKelvey and Palfrey (1998) defines a sequential game analogue of the QRE model where players are playing error prone strategies. As discussed above, these theories do not model the specific form of bounded rationality that we model in LFE, i.e., limited foresight. The distinction is important because limited foresight generates particular patterns (Rampal (2017)). Three examples of such data patterns are: first, failure to play dominant strategies in a sequential game; second, a significant decline in the proportion of dominated choices towards the end-stages of the game; third, evolution of players' beliefs (hence strategies) about the opponent's expertise within a play of the game based on the opponent's past moves in the game.

The reputation effects in our model are similar to the crazy type literature started by Kreps, Wilson, Milgrom and Roberts in 1982, yet there are important differences. Unlike this paper, their crazy types' behavior is exogenous, and their crazy types (whose counterparts in the LFE model would be the player-types with low foresight) are given no incentive to discover whom they are playing against.

2 Model

The model that we define here seeks to capture the scenario where players are playing what "seems to be" a finite sequential game with perfect information. Specifically, all prior actions taken in the game are observed by every player at his move, but every player has a particular level-of-expertise/experience in the game and every player is uncertain about each opponent's level-of-expertise/experience in the game. In particular, we will focus on the case where this level of expertise/experience translates into a level of foresight. That is, we will model the scenario where every player can have one of various possible levels of foresight and every player is uncertain about the level of foresight of each of his opponents.[10] The foresight level of a player is defined as the number of subsequent stages that a player can observe from any given move. This model is in the paradigm of selfish utility maximization. That is, we will not model any form of uncertainty about players' utilities due to other-regarding preferences or game-specific types, etcetra. We will assume that all relevant information about payoffs

[10] We don't model *how* or *why* the expertise/experience translates into a particular level of foresight. This is left for future research.

is captured by the payoffs of the finite sequential game with "seemingly" perfect information that the players are playing.

To model limited foresight and uncertainty about the opponents' foresight, we start with the game that the players "seem to be" playing, i.e. a finite sequential game with perfect information called Γ_0. As an example, think of Γ_0 as the game with perfect information that the experimenter sets up for a set of players to play. We map Γ_0 to the game that is "actually" being played, i.e., a standard Bayesian game with imperfect information called Γ. The construction of Γ from Γ_0 is done by introducing a set of possible types corresponding to each player in Γ_0, where each type denotes a particular level of foresight. For example, in the experiment in Rampal (2017), at the beginning of every round players were randomly paired up to play what "seemed to be" a finite sequential game with perfect information: the "race game" (the Γ_0 in our model). However, subjects were aware that about half the subjects were experienced in the race game, while the other half were inexperienced. Before the game began, each player was uncertain about the experience-level of his/her randomly paired opponent. So to model this uncertainty one must transform the race game (Γ_0) to a "race game with uncertainty about the opponent's experience-level" (Γ).[11]

In Γ, each player-type is uncertain about each opponent's level of foresight or, equivalently, each opponent's type. The Limited Foresight Equilibrium (henceforth LFE) provides a strategy and a belief profile for Γ. However, we cannot solve for the LFE actions and beliefs of player-types who have limited foresight using Γ because player-types with limited foresight cannot observe Γ at their moves. Therefore we will consider appropriately curtailed versions of Γ, which is what the limited-foresight player-types observe, to solve for the LFE strategy and belief profile of Γ. In subsection 2.1 we define the underlying sequential game with perfect information Γ_0. In 2.2 we construct Γ, the sequential game with imperfect information, from Γ_0. In 2.3 we construct the curtailed versions of Γ which are observed by player-types possessing limited foresight.

[11]Rampal (2017) found that beliefs about the opponent's experience level were indeed significant in determining optimal behavior.

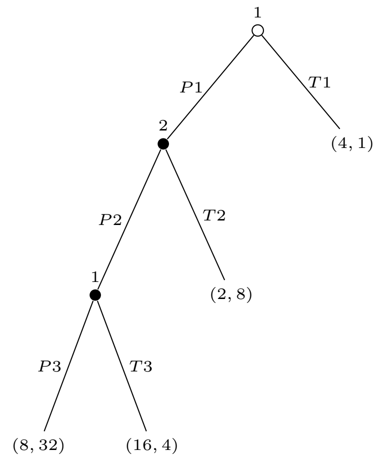

Figure 1. An Underlying Sequential Game with Perfect Information: Three-Staged Centipede Game
Notes. The LFE model is applicable to all finite sequential games with perfect information. We use the 3-staged Centipede Game to illustrate LFE because it is simple to draw as it has only 4 terminal histories despite being a 3-staged game.

2.1 The Underlying Sequential Game with Perfect Information

We use the standard notation from Osborne and Rubinstein (1994), with minor modifications, to define a sequential game with perfect information and perfect recall called Γ_0.[12] In particular, Γ_0 is defined as a collection of the following components.

- A finite set of players N_0.
- A set H_0 of finite sequences of actions or histories such that:
  - The empty sequence $\emptyset$ is a member of H_0.
  - If a sequence of actions $h_0 = (a_0^k)_{k=1,\ldots,K} \in H_0$ and $L < K$ holds, then $(a_0^k)_{k=1,\ldots,L} \in H_0$. The k^{th} action, a_0^k, is said to be taken at the k^{th} stage of the game. The set of terminal histories, denoted as Z_0, is defined as the set of histories $(a_0^k)_{k=1,\ldots,K} \in H_0$ such that there is no $(K+1)$ such that $(a_0^k)_{k=1,\ldots,K+1} \in H_0$.

[12]We don't specify the conditions for perfect recall.

- A set of possible actions in the game, A_0, and an action correspondence $A_0(.)$ which maps $h_0 \in H_0$ to a set $A_0(h_0) \equiv \{a_0 \in A_0 \ : \ (h_0, a_0) \in H_0\}$.

- A function $P_0(.)$, called the player function that assigns a player $P_0(h_0)$ to each history h_0. That is, $P_0(.)$ maps each element of H_0 to an element in N_0.

- For each player $i \in N_0$, a Bernoulli utility function u_i which maps terminal histories to real numbers. That is, u_i maps a terminal history $z_0 \in Z_0$ to a payoff $u_i(z_0) \in \mathcal{R}$.

Thus, Γ_0 is defined as $\{N_0, H_0, P_0, A_0, \{u_i\}_{i \in N_0}\}$. Let the number of stages in Γ_0 be S.[13] Consider the three-staged centipede game in Figure 1 as an example of an underlying sequential game with perfect information.

2.2 Modeling Uncertainty: From Perfect Information to Imperfect Information

We now construct a sequential game with imperfect information called Γ from the sequential game with perfect information, Γ_0. The only form of imperfection in information we allow in Γ relative to Γ_0 comes from the aim to model a scenario where each player $i \in N_0$ has several possible types and each player's type is his private information. In particular except Nature's initial move, which determines the probability distribution over possible combinations of players'-types playing each other, all other prior actions will be known at each move. For example, consider the case where $N_0 = \{\text{Ann, Bob}\}$ are playing Γ_0 in the underlying game. To define Γ, in the very first stage of Γ, we will introduce Nature's move which specifies the probability distribution over the possible combinations of types of the players in N_0. Roughly speaking, after Nature moves, Ann and Bob will play Γ_0 knowing his/her own type, but without knowing his/her opponent's type. For example, suppose Γ_0 is tic-tac-toe. Suppose that Nature specifies that "type t_{Ann} of Ann has a 30 percent chance of occurring and type t'_{Ann} of Ann has a 70 percent chance of occurring. Independently, type

[13] Formally, $S \equiv max\{K\colon\ (a_0^k)_{k=1,\dots,K} \in H_0\}$.

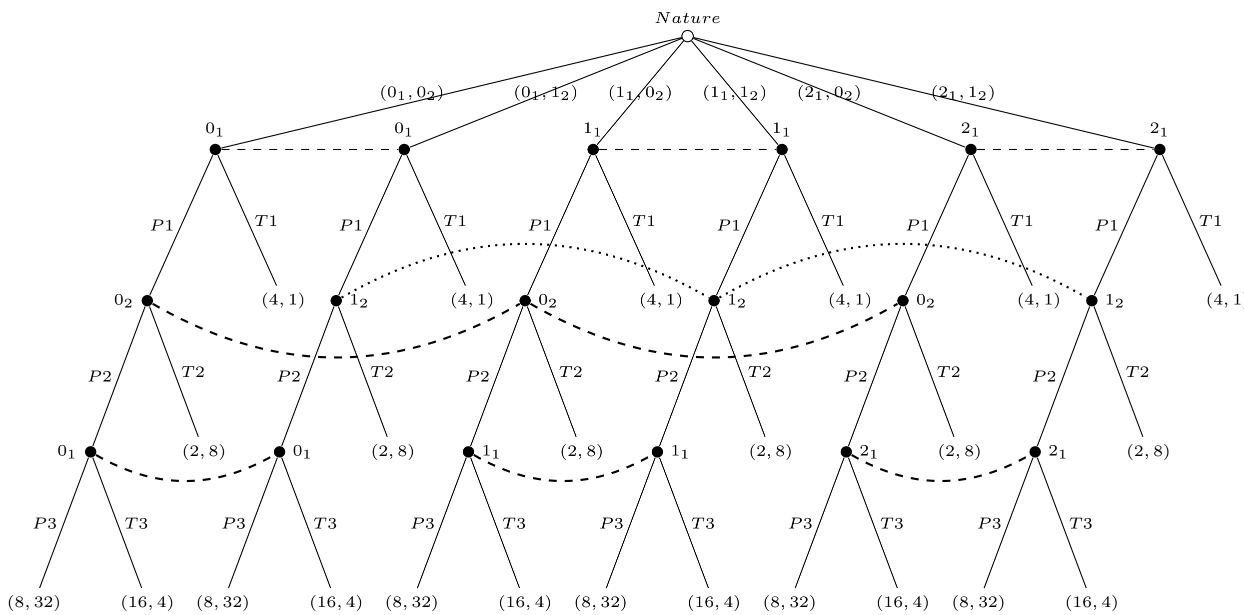

Figure 2. The Three-Staged Centipede Game Converted to an Interaction Game
Notes. The figure shows the conversion of Γ_0 (the Centipede game depicted in Figure 1) into an Interaction Game, Γ, depicted here. As there are 3 stages in the Centipede game in Figure 1, so there are 3 types of player-1 possible, $\{0_1,\ 1_1,\ 2_1\}$, and two types of player-2 possible, $\{0_2,\ 1_2\}$. For each combination of player-1's type and player-2's type (who could be playing Γ_0 with each other) we redraw Γ_0 to generate Γ. We construct the information sets so that each player-1 type, at each of his moves, observes the sequence of prior actions played, but doesn't know which player-2 type (0_2 or 1_2) he is playing against. Similarly, each player-2 type, at each of her moves, observes the sequence of prior actions played, but doesn't know which player-1 type (0_1, 1_1 or 2_1) she is playing against. Suppose ρ, the common-knowledge prior probability distribution on players'-types is $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ on $\{0_1, 1_1, 2_1\}$, and independently, $(\frac{1}{10}, \frac{9}{10})$ on $\{0_2, 1_2\}$. That is, let $\rho(0_1, 0_2) = \rho(1_1, 0_2) = \rho(2_1, 0_2) = \frac{1}{30}$ and $\rho(0_1, 1_2) = \rho(1_1, 1_2) = \rho(2_1, 1_2) = \frac{9}{30}$.

t_{Bob} of Bob has a 60 percent chance of occurring and type t'_{Bob} of Bob has a 40 percent chance of occurring." Then in Γ, after this Nature's move, Ann will play tic-tac-toe with Bob knowing her type, t_{Ann} or t'_{Ann}, whichever it may be, but without knowing if Bob's type is t_{Bob} or t'_{Bob}. As an illustrative example of what we are trying to do, consider the conversion of the Centipede game in Figure 1 to an Interaction Game depicted in Figure 2.

We know proceed to formally defining the construction of Γ from Γ_0. The Interaction Game Γ (which is a finite sequential game with imperfect information), constructed from Γ_0, has the following components.

- A finite set of player-types: N. To construct N from N_0, each $i \in N_0$ generates a corresponding set of i-types in Γ called T_i. Let s_i denote the first stage of Γ_0 at which i moves,[14] then the set of possible types of player i in Γ is given by $T_i \equiv \{0_i, 1_i, .., (S-s_i)_i\}$. So $N = \bigcup_{i \in N_0} T_i$. Further, a particular combination of player-types of Γ is $t \in T = \times_{i \in N_0} T_i$. A player-type, for example 1_i, should be interpreted as follows. 1_i is type 1 of player $i \in N_0$. 1_i is that type of player i who has the foresight-level of 1, i.e., given his move in Γ, 1_i can observe one more stage after his move. We model limited foresight more comprehensively in the next subsection.

- A set of sequences H, and a set of terminal sequences Z. The sets H and Z are generated from H_0 using a set valued mapping Seq: $H_0 \implies H$. The set $Seq(h_0)$, a subset of H, contains elements of the form $(t,\ h_0)$. That is, the Seq: $H_0 \implies H$ correspondence replicates each sequence h_0 in Γ_0 for each possible player-type combination t in Γ. So $Seq((a_0^k)_{k=1,...,K}) = \{(t,\ (a_0^k)_{k=1,...,K})\colon\ t \in T\}$. Thus, each sequence $h_0 = (a_0^k)_{k=1,...,K} \in H_0$ corresponds to a set of sequences in H, and H is defined as $\bigcup_{h_0 \in H_0} Seq(h_0)$.

- A player function P, which maps each element of $H \backslash Z$ to an element in N. The function P has the following properties.

  - $P(\emptyset)$ =Nature.[15]
  - If player i moved after $h_0 = (a_0^k)_{k=1,...,K} \in H_0$ in the underlying game, then $t_i \in T_i$ moves after $((t_i, t_{-i}),\ h_0) \in H$ for all $t_{-i} \in T_{-i}$. That is, consider an arbitrary $((t_i, t_{-i}),\ h_0) \in H$. If $P_0(h_0) = i$, then $P(((t_i, t_{-i}),\ h_0)) = t_i$.

[14] Formally, if i moves at $\emptyset$ in Γ_0, then $s_i = 1$. Otherwise, $s_i \equiv min\{K\colon (a_0^k)_{k=1,...,K-1} \in H_0$ and $P_0((a_0^k)_{k=1,...,K-1}) = i\}$

[15] We say that the Nature's $\emptyset$ history action is taken at the 0^{th} stage.

- Nature's move which specifies a probability distribution on T. This distribution, denoted by ρ, is assumed to be common knowledge.[16] Further, $\rho(t) \in [0,1]$ for all $t \in T$, and $\sum_{t \in T} \rho(t) = 1$ hold.

- A set of possible actions in the game, A, and an action correspondence $A(.)$.

  - $A(.)$ maps $h \in H$ to a set $A(h) \equiv \{a \in A \ : \ (h, a) \in H\}$.
  - The set of possible actions, or action set, after a sequence $h \in H$ is the same as the action set after the corresponding $h_0 \in H_0$ that generated h. Formally, consider an arbitrary $h \in H$ such that $h \neq \emptyset$. Let $h = (t,\ h_0)$ (or, put another way, $h \in Seq(h_0)$). Then $A((t,\ h_0)) = A_0(h_0)$.

- For each player-type t_i, a partition $\mathcal{I}(t_i)$ of the set of histories/sequences of H where t_i moves, i.e., a partition $\mathcal{I}(t_i)$ of $\{h \in H\colon\ P(h) = t_i\}$. An information set of t_i in $\mathcal{I}(t_i)$ is denoted as $I(t_i)$. The information sets in $\mathcal{I}(t_i)$ obey the usual restriction that the actions available from and the player-type moving at all histories of an information set must be the same.[17] The construction of Γ from Γ_0, described above, gives us more structure. Consider an arbitrary history h_0 of Γ_0. Suppose i is the player moving after this sequence of actions, h_0, in the underlying game. That is, let $P_0(h_0) = i$. Then h_0 will map to $Seq(h_0) = \{(t,\ h_0)\colon\ t \in T\}$ in Γ. The set $\{(t,\ h_0)\colon\ t \in T\}$ will be subdivided into $|T_i|$ (the cardinality of the set T_i) information sets in Γ, one information set for each $t_i \in T_i$. Further, for each t_i, if t_i's information set $I(t_i)$ is a subset of $Seq(h_0)$, that is $I(t_i)$ is generated from h_0 in Γ_0, then it must be the case that $I(t_i) = \{((t_i,\ t_{-i}),\ h_0)\colon\ t_{-i} \in T_{-i}\}$. So, at each such information set of Γ the player-type moving there, say t_i, is aware about all prior actions (given by h_0); but t_i

[16] The common-knowledge-prior assumption helps provide a tractable structure to a lot of the following analysis. Even with this assumption, we will see that the beliefs of players with different foresight-levels evolve differently upon observing the same sequence of prior actions. We discuss this assumption in more detail in the last section, which defines the LFE for the case of subjective prior beliefs, that is, for the case where prior beliefs are allowed to differ across player-types.

[17] Formally, $P(h) = t_i\ \forall h \in I(t_i)$, and $A(h) = A(h')\ \forall h,\ h' \in I(t_i)$.

is uncertain about which combination of opponents' types (t_{-i}) he is playing against. As $\bigcup_{h_0 \in H_0} Seq(h_0) = H$ holds, all the information sets of Γ obey this structure.

- For each player-type $t_i \in N$, a Bernoulli utility function u_{t_i} which maps terminal histories Z to real numbers. These u_{t_i} functions also obey additional structure. In particular, for each $z \in Z$, if $z \in Seq(z_0)$, then the utility derived by an arbitrary player-type t_i at z, denoted as $u_{t_i}(z)$, is equal to the utility derived by i at the corresponding $z_0 \in Z_0$. That is, $u_{t_i}(z) = u_i(Seq^{-1}(z))$, $\forall\ t_i \in T_i$, $\forall\ i \in N_0$, and for all $z \in Z$.

Thus, the Interaction Game $\Gamma = \{N,\ H,\ \{\mathcal{I}(t_i)\}_{t_i \in N},\ P,\ A, \{u_{t_i}\}_{t_i \in N}, \rho\}$, corresponding to the underlying sequential game with perfect information Γ_0, is defined by its construction using the $Seq(.)$ correspondence.

2.3 Modeling Both Limited-Foresight and Uncertainty: Curtailed Games

The Limited Foresight Equilibrium (LFE) provides an outcome prediction for the Interaction Game Γ by specifying an "equilibrium" strategy profile and the associated belief profile for it. We put "equilibrium" in quotes because LFE cannot be solved using the Interaction Game. We can't use a solution concept directly on the Interaction Game because limited-foresight player-types may not observe the Interaction Game at their information sets, hence they cannot optimize based on the Interaction Game. At each of their moves, limited-foresight player-types optimize based on a curtailed version of the Interaction Game that *they* observe from that move given their limited foresight. That is, player-types use their move specific curtailed version of the Interaction Game to optimize. These curtailed versions of the Interaction Game are said to be the *Curtailed Games* generated from the Interaction Game.

As the name suggests, a *Curtailed Game* is defined by curtailing the Interaction Game at a particular stage. Consider an Interaction Game with S-stages: Γ. An n-staged Curtailed Game constructed from Γ will be labeled as $CG(n)$. Intuitively, $CG(n)$ will be constructed by curtailing the Interaction Game after the n^{th} stage actions. Let

$$CG(n) = \{N,\ H^n,\ \{\mathcal{I}^n(t_i)\}_{t_i \in N},\ P^n,\ A^n, \{u^n_{t_i}\}_{t_i \in N}, \rho\}.$$

The components of $CG(n)$ are defined as follows.

- $CG(n)$ is an exact replica of Γ until (and including) stage $(n-1)$. The player set N of $CG(n)$ is the same as the player set N of Γ. H^n, the set of histories of $CG(n)$, is defined as $(t,\ (a^k_0)_{k=1,\dots,K}) \in H$ such that $K \leq n$ and $t \in T$. Let Z^n be the set of terminal histories of $CG(n)$ (defined in the next bullet point). The set of non-terminal histories of $CG(n)$, $(H^n \backslash Z^n)$, is partitioned into information sets, $\{\mathcal{I}^n(t_i)\}_{t_i \in N}$, exactly as $(H \backslash Z)$ is partitioned into $\{\mathcal{I}(t_i)\}_{t_i \in N}$. The player function $P^n(h) = P(h)$, and the action function $A^n(h) = A(h)$, for all $h \in (H^n \backslash Z^n) \bigcap (H \backslash Z)$. Nature's common-knowledge prior distribution over T, given by ρ in Γ, is the same in every $CG(n)$ and it remains common knowledge in every $CG(n)$.[18]

- The set of terminal histories of $CG(n)$, denoted as Z^n, contains two kinds of terminal histories. First, the terminal histories of Γ which end at or before an n^{th} stage action. These terminal histories are collected in $Z^n(1) \subset Z^n$. Formally $Z^n(1) = \{(t,\ (a^k_0)_{k=1,\dots,K}) \in Z\colon\ K \leq n \text{ and } t \in T\}$. Second, those sequences/histories of the Interaction Game, $(t,\ (a^k_0)_{k=1,\dots,K})$, with $K > n$, which are curtailed at $(t,\ (a^k_0)_{k=1,\dots,n})$, converted to terminal histories and collected in $Z^n(2) \subset Z^n$ for the construction of $CG(n)$. Formally, $Z^n(2) = \{(t,\ (a^k_0)_{k=1,\dots,n}) \in (H \backslash Z)\}$. The set of terminal histories

[18]The common-knowledge prior distribution assumption helps in reducing the number of possible different ways the various limited-foresight players can observe a curtailed version of the Interaction Game. If every player-type has a different subjective prior belief over opponents' types, then we must construct a move-specific Curtailed Game for each individual player-type. The common-knowledge prior distribution lets us consider only S possible curtailed versions of Γ for the purpose of solving for the strategies and beliefs of all the player-types. The benefit of this common-knowledge prior distribution assumption will become clearer when we define LFE (next section). One method for weakening this assumption will be specified when we model subjective prior beliefs (last section).

of $CG(n)$ is $Z^n = Z^n(1) \bigcup Z^n(2)$.

- *Payoffs.* For those terminal histories of $CG(n)$ which are also the terminal histories of Γ, the payoffs of each player-type remain the same. That is, for all $z^n \in Z^n(1)$, $u^n_{t_i}(z^n) = u_{t_i}(z^n)$. The "controversial" choice that must be made in curtailing the Interaction Game is that "what is the payoff profile associated with those terminal histories of $CG(n)$ which are not terminal histories of Γ?" Any payoff numbers placed at such "synthetic" terminal histories in $Z^n(2)$ of $CG(n)$ will have to follow *some* rule. We use the $[(min + max) \div 2]$ "curtailment" rule of Mantovani (2014).[19] Consider the terminal sequence $h \in Z^n(2)$ of $CG(n)$. The $[(min + max) \div 2]$ rule implies that each player-type's payoff after h in $CG(n)$ is the average of the minimum and the maximum that that player-type could achieve in Γ following all possible terminal action sequences after h. That is, for each $h \in Z^n(2)$, let $Z(h)$ be a subset of Z such that for any arbitrary terminal history z in $Z(h)$, the actions in the first n stages of z are played as specified in h. Formally, let $Z(h) = \{z \in Z\colon\ z = (h, (a^k_0)_{k=n+1,\dots,S})\}$ (where S is the number of stages in Γ). Then, for each $t_i \in N$, $u^n_{t_i}(.)$ is defined over $Z^n(2)$ as follows.

$$u^n_{t_i}(h) = \frac{min\{u_{t_i}(z)\colon\ z \in Z(h)\} + max\{u_{t_i}(z)\colon\ z \in Z(h)\}}{2} \text{ for all } h \in Z^n(2) \qquad (1)$$

As an example of a Curtailed Game, consider the one-staged Curtailed Game, $CG(1)$ depicted in Figure 3, constructed from the Interaction Game in Figure 2.

[19] Shaowei Ke (2017) discusses these rules axiomatically. Our contribution is to model limited foresight *with* uncertainty *and* updating about the opponent's foresight within a play of the game. Thus, we stick to a simple curtailment rule and model these features with this simple rule. In an older version of this paper we used a "mean of stage-wise means" rule explained there, which didn't change any of the results that follow. The interested reader can access the paper by an emailed request at jeevantr@iima.ac.in.

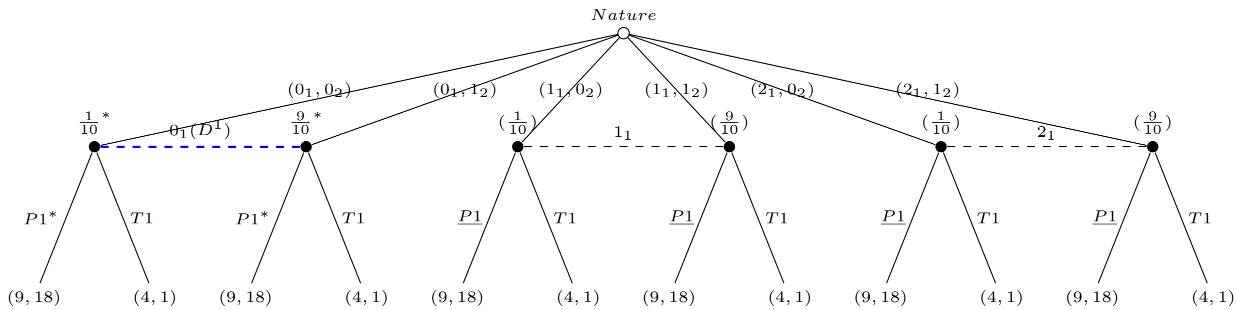

Figure 3. Curtailed Game (1)
Notes. The figure shows the conversion of Γ, the Interaction Game depicted in Figure 2, into its shortest Curtailed Game, $CG(1)$ (short for Curtailed Game (1)), depicted here. $CG(1)$ is identical to Γ in all respects except that $CG(1)$ ends after the first stage action. If any type of player-1 chooses $T1$ (*take* in stage 1) in the first stage then the associated $CG(1)$ payoff profile is $(4, 1)$, identical to the Interaction Game. However, we need to specify a payoff profile for the case where a player-1 type chooses $P1$ (*pass* in stage 1). As shown in the construction of $CG(1)$, we curtail the Interaction Game after the action $P1$ and use the $\frac{min+max}{2}$ rule for payoffs. For example, after playing $P1$, the maximum a player-1 type can get in Γ is 16 and the minimum he can get is 2, thus, his payoff from choosing $P1$ in $CG(1)$ is $\frac{min+max}{2} = 9$. We mark $0_1's$ first-stage information set as D^1. This is because $CG(1)$ is exactly what 0_1 observes at his first-stage information set. Thus, $CG(1)$ is *decisive* for 0_1 at stage-1. $CG(1)$ represents the payoff profiles observed by 0_1 at stage-1, thus any "reasonable" notion of optimality should specify that 0_1 chooses $P1$ (asterisked) at his stage-1 information set and has the belief $(\frac{1}{10}, \frac{9}{10})$ (asterisked) over $(0_2, 1_2)$ as per the prior distribution specified in Γ (Notes of Figure 2). These assertions and the reason behind underlining the action $P1$ for 1_1 and 2_1 at their respective information sets and bracketing their beliefs will be made precise when we define LFE.

3 Limited Foresight Equilibrium

We now proceed to defining the Limited Foresight Equilibrium (LFE). The first step in this direction is defining *total foresight.* The total foresight of a limited-foresight type determines the Curtailed Game that that limited-foresight type observes. A limited-foresight type's *total foresight* is defined as the sum of (i) the stage number that the limited-foresight type is moving at, and (ii) the level of foresight of that limited-foresight type. If this sum is greater than the total number of stages in the Interaction Game, say S, we say that the total foresight is S. We make this precise in Definition 1 below. In what follows, let the foresight-level of player-type t_i be denoted as t_i itself. For example, the player-type 3_i has a foresight-level of 3. We denote the foresight-level of 3_i as 3_i, it is understood that the foresight level is actually 3.[20] The definition of total foresight is as follows.

Definition 1 *(total foresight)*: consider a sequence $h = (t,\ (a_0^k)_{k=1,\ldots,s-1}) \in H$ of an S-staged Interaction Game. Suppose the player-type $P(h) = t_i$ moves at the s^{th}-staged sequence h. Let $P(h) = t_i$. The *total foresight* of player-type t_i at stage s is $min\{(t_i + s), S\}$.

As an example of the application of Definition 1, consider the 3-staged Centipede Game example in Figures 1-5, the total foresight of player-type 0_1 (foresight-level of 0) at stage-one is equal to 1 and this total foresight is captured by $CG(1)$ (Figure 3). The total foresight of 0_2 at stage-two and 1_1 at stage-one is 2, and this is captured in $CG(2)$ (Figure 4). The total foresight of 0_1, 1_1, and 2_1 at stage-three, 1_2 at stage-two, and 2_1 at stage-one is 3 (the total number of stages), and this is captured in $CG(3)$ (Figure 5).

The next definition, makes the phrase "the Curtailed Game that a limited-foresight type observes at a certain move" precise. In particular, we define *decisive information sets* of a Curtailed Game and the *decisive Curtailed Game* for an information set.

Definition 2 *(decisive information sets and decisive Curtailed Games)*: Let Γ be an S-staged Interaction Game. Consider an n-staged Curtailed Game, $CG(n)$, constructed from Γ. $CG(n)$ is said to be *decisive* for the information sets D^n of Γ, iff for all $I \in D^n$, the player-type moving at I, $P(I)$, has the total foresight of n at I. The information sets in D^n are said to be *decisive* for $CG(n)$.

Definition 2 implies that the player-types moving at the *decisive information sets* of a Curtailed Game, say $CG(n)$, observe exactly $CG(n)$ at these information sets due to their respective foresight-level and the stage number of their respective decisive information set. Thus, they use $CG(n)$ to optimize, so $CG(n)$ is decisive for these information sets. In our

[20] So if 3_i is moving at stage-4, his total foresight is $3_i + 4 = 7$. This abuse of notation helps simplify the notation.

Centipede game example, $CG(1)$ is decisive for 0_1 at his stage-1 information set, $CG(2)$ is decisive for 0_2 at stage-2 and 1_1 at stage-1, and $CG(3)$ is decisive for all other information sets.

In defining the LFE and discussing its features, it will be useful to establish a hierarchy among player-types based on their total foresight. The next definition establishes higher and lower types relative to a given player-type with a given total foresight.

Definition 3 *(higher types and lower types)*: Consider a player-type t_i moving at an information set $I(t_i)$ of the Interaction Game Γ. Suppose t_i has total foresight equal to n at $I(t_i)$. So $CG(n)$ is decisive for t_i at $I(t_i)$. An arbitrary player-type t_j, moving at an information set $I(t_j)$ of $CG(n)$, is considered a higher (respectively lower) type than t_i at $I(t_i)$, if t_j's total foresight at $I(t_j)$ is weakly greater than n (strictly lower than n).

Consider the following examples of higher and lower types. Consider the 3-staged Centipede Game example in Figures 1-5. For 0_1 at stage-one (total foresight of 1 (Figure 3)), both 1_1 and 2_1 at their respective information sets in stage-one are higher types. For 1_1 at stage-one and 0_2 at stage-two (total foresights of 2 (Figure 4)), only 0_1 at stage-1 is a lower type, and player-types at other information sets in $CG(2)$ are higher types. For 2_1 at stage-one, 1_2 at stage-two, and all player-1-types at stage-three (total foresights of 3 (Figure 5)), 0_1 and 1_1 at stage-1, and 0_2 at stage-2 are lower types (because their total foresights are equal to 2), and player-types at other information sets (including 0_1 at stage-3) are higher types.

A Rule-of-Thumb Definition of LFE

The definition of LFE (we state the precise definition later) boils down to three **rules-of-thumb**. Suppose we have already constructed an Interaction Game Γ from a game with perfect information using ρ, the common knowledge prior distribution over foresight-levels across players. Suppose now we are trying to solve for the LFE action of the limited-foresight player-type t_i moving at some information set $I(t_i)$. Let n be t_i's total foresight at $I(t_i)$. So t_i observes $CG(n)$ at $I(t_i)$. The three rules-of-thumb for solving for t_i's LFE action and belief at $I(t_i)$ are:

(a) t_i knows the LFE actions of lower types. He considers these actions fixed as Nature's moves in $CG(n)$.

(b) t_i assumes that higher types, and t_i himself, choose a strategy for $CG(n)$. Each higher type's chosen strategy in the strategy profile so constructed must be a sequentially rational best response in $CG(n)$ given the "fixed" LFE actions of lower types (as per rule (a)), the rest of the strategy profile, and beliefs.

(c) t_i's beliefs, and the beliefs of all other player-types in $CG(n)$ are calculated using the Bayes' rule (whenever possible), given Nature's moves (as per rule (a)), the strategy profile described in (b), and ρ, the common prior.

All of these rules are captured in the following definition: Formally, we solve for t_i's LFE action and belief at $I(t_i)$ by solving for the Sequential Equilibrium (SE) of $CG(n)$, after taking the LFE actions in (a) given as Nature's moves. As we first apply these rules at $CG(1)$, then proceed to $CG(2)$, and then step-wise to longer Curtailed Games, rule (a) is well defined at each step. Further, in the next section, Proposition 2 will show that, under certain conditions, calculating the SE of each Curtailed Game, say $CG(n)$, is straightforward as different types of each player can be treated as the same type if their total foresight is n or more (and strategies of player-types with total foresight strictly less than n is given by rule (a)).

As an example of these rules, consider the 3-staged Centipede Game again. Consider Figure 4, which depicts the method for solving for the LFE action and belief of 1_1 at stage-one. As the total foresight of 1_1 at stage-one is 2, we use $CG(2)$ to solve for 1_1's LFE action at his stage-one information set. Rule (a) implies that at 0_1's stage-one information set, "Nature" replaces 0_1 as the player moving there. Further, 0_1's LFE action ($P1$) at his stage-one information set is considered as Nature's move and marked as $P1^S$, where the superscript denotes "solved." Thus, rule (a) implies "modifying" $CG(2)$ and constructing the *Modified Curtailed Game(2)* ($MCG(2)$ for short) from $CG(2)$. Figure 4 depicts $MCG(2)$. We will formally define a MCG in the next sub-section. Rule (b) implies that 1_1, 2_1, 0_2 and 1_2, moving at their respective information sets in $CG(2)$ are higher types compared to 1_1 at stage-one; thus, they must choose actions which are sequentially rational in $CG(2)$ given Nature's (acting on behalf of 0_1) stage-one action, i.e. $P1^S$, the actions chosen by others, and their beliefs at their respective information sets. Rule (c) implies that the beliefs used in rule (b) must be calculated using the prior distribution, the strategies specified in rules (a) and (b), and Bayes' rule. All three rules, (a), (b) and (c) are satisfied if we simply solve for the SE of $CG(2)$ after taking 0_1's stage-one action, $P1^S$, as Nature's move. As Figure 4 shows, the LFE action of 1_1 at stage-one is $P1$ (asterisked) and his beliefs (asterisked) over player-2's types are the same as the prior distribution over player-2's types.

It is worth noting that due to the assumption that the prior distribution over players' foresight-levels is common knowledge, 0_2 at stage-two observes exactly the Curtailed Game observed by 1_1 at stage-one. Thus, we can use the same Curtailed Game, $CG(2)$, to solve for 0_2's stage-two LFE action.[21] So in solving the SE of $CG(2)$, we have also solved for the LFE action and beliefs of 0_2 at his stage-two information set: 0_2 chooses $P2$ in stage-two and his beliefs over player-1's types are the same as the prior distribution over player-1's types. The next step would be to consider $CG(3)$, replace the LFE actions for the decisive information sets of $CG(1)$ and $CG(2)$ as Nature's moves in $CG(3)$ and solve for the SE of $CG(3)$. Proceeding from the shortest Curtailed Game to the Interaction Game as above gives us the LFE for an Interaction Game. Figure 5, completes the construction of LFE for the 3-staged Centipede Game example.

[21] This would not be the case if 0_2's prior belief were different from 1_1's prior belief. In that case, in order to consider the CG that 0_2 uses to optimize, we would have to construct a different $CG(2)$, one with Nature's initial distribution over foresight levels as per 0_2's prior beliefs. We retain the common prior assumption until the last section where we discuss subjective prior beliefs. Even with the common prior assumption, we will see that the beliefs of players with different foresight-levels evolve differently upon observing the same sequence of prior actions.

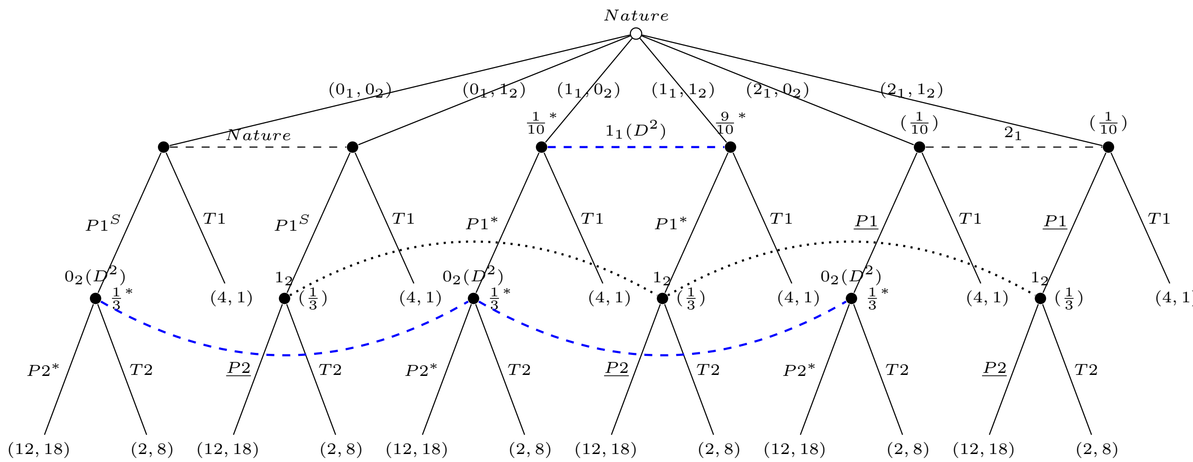

Figure 4. Modified Curtailed Game (2)

Notes. From $CG(1)$, depicted in Figure 3, we know that in LFE, 0_1 chooses $P1$ (*pass* at stage-1) at his first-stage information set, D^1, for which $CG(1)$ is decisive. $MCG(2)$ is constructed from $CG(2)$ in two steps: First, the player function of $MCG(2)$ specifies *Nature*, instead of 0_1, as the player moving at D^1; second, $0_1's$ LFE action at D^1 is considered as common-knowledge Nature's move and marked with the superscript S, which denotes "solved" (so $P1$ becomes $P1^S$). $MCG(2)$ is identical to Γ in all respects except that (i) $MCG(2)$ curtails Γ after the second-stage action and (ii) in $MCG(2)$, we take $0_1's$ LFE action at D^1 as Nature's move. As *both* 1_1 at stage-1 and 0_2 at stage-2 observe two stages of the Interaction Game, $MCG(2)$ is decisive for *both* these information sets, denoted by D^2. Note that it is strictly better for all player-2 types to choose $P2$, irrespective of beliefs. Thus, in any SE of $MCG(2)$, and so in LFE, 1_1 chooses $P1$ (asterisked) in stage-1 and 0_2 chooses $P2$ (asterisked) in stage-2, irrespective of their beliefs. Consistent with this SE of $MCG(2)$, the LFE beliefs (asterisked) of 1_1 at stage-1 and 0_2 at stage-2 are their priors over the opponent's type given from Γ (Figure 2 notes). Note that we mark the SE actions and beliefs at even the non-decisive information sets of $MCG(2)$ by underlining them and bracketing them, respectively, as we did in $CG(1)$ (Figure 3). These beliefs are the second order beliefs of 1_1 at stage-1 and 0_2 at stage-2. These SE actions and beliefs at non-decisive information sets are calculated because they are needed to calculate the LFE actions and beliefs at decisive information sets (in general, but not in this example).

Note that in solving for the Sequential Equilibrium of a Curtailed Game, we are also solving for the SE actions and beliefs of player-types at information sets other than the decisive information sets of that Curtailed Game. For example, in solving for the SE of $CG(2)$, we also solved for the SE actions and beliefs of 2_1 at his stage-one information set. We do not consider the SE actions at the "non-decisive" information sets of a Curtailed Game as LFE actions. These actions, for example 2_1's action of $P1$, are typically needed to calculate the LFE actions and beliefs at the decisive information sets of the Curtailed Game (in most Curtailed Games, but not this $CG(2)$ example). Using the SE for $CG(2)$ implies that at their stage-two and stage-one information sets, respectively, 0_2 and 1_1 think that other player-types will play sequentially rational strategies for $CG(2)$ and that they will have beliefs consistent with the strategy profile in $CG(2)$.[22] Note that the SE beliefs of player-types at these "non-decisive" information sets of a Curtailed Game are the second-order beliefs of the players at the decisive information sets. In our example, in the SE of $CG(2)$, the SE beliefs of 2_1 at his stage-one information set in $CG(2)$ are the second-order beliefs of 0_2 and 1_1. That is, the SE beliefs of 2_1 represent what 0_2 and 1_1 (at stage-two and stage-one respectively) believe are the beliefs of 2_1 at stage-one. These second-order beliefs are important in that they help determine 2_1's stage-one SE action, which in turn determines the SE (and LFE) actions of 0_2 and 1_1 at their stage-two and stage-one information sets, respectively. In what follows, we assume that the belief hierarchy terminates at these second-order beliefs.

Suppose a player-type t_i has total foresight n at an information set $I(t_i)$. Further suppose that the LFE strategy profile is such that at $I(t_i)$, t_i is facing a higher type with probability equal to one, and t_i knows this. One might ask, why should t_i still use the $CG(n)$ he observes to try and understand his opponent's choice? Our answer to this question is that $CG(n)$ represents the limit of the foresight of t_i at $I(t_i)$, and he uses this limit to try and reason

[22] Selecting a common SE for $CG(2)$ to represent what both 0_2 and 1_1 think at their respective information sets is a significant assumption when there are multiple SE of the Curtailed Game. We persist with this definition for simplicity and in keeping with the notion of an equilibrium.

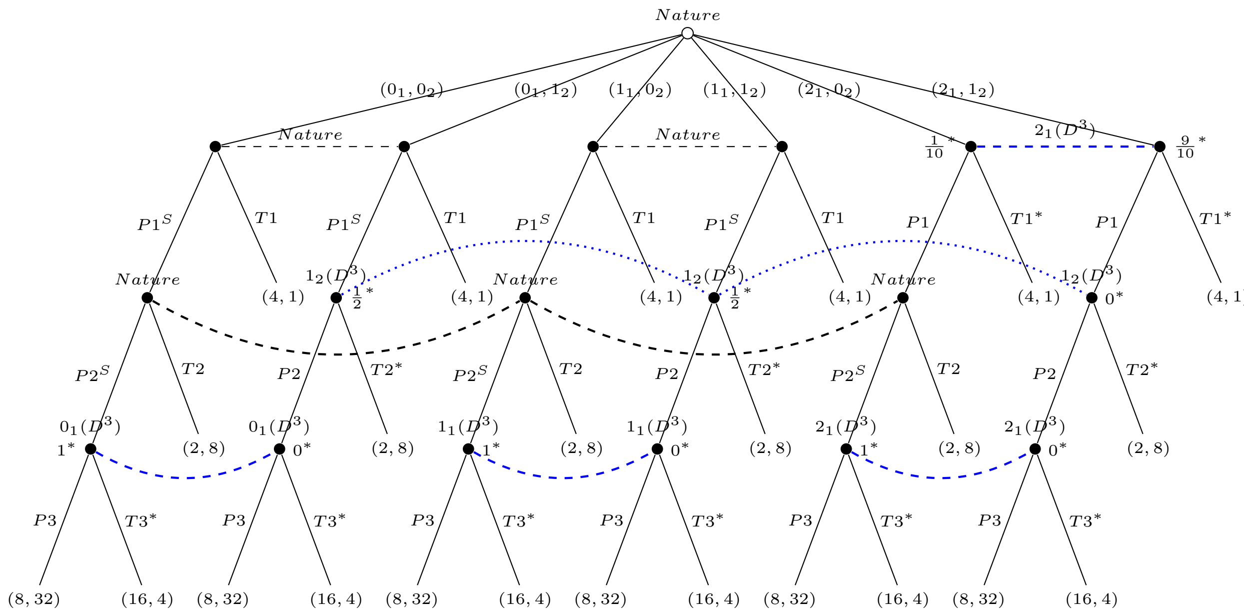

Figure 5. Modified Curtailed Game (3): Solving for the LFE.

Notes. From Figure 3 ($CG(1)$) and Figure 4 ($MCG(2)$) we know the LFE actions and beliefs of 0_1 and 1_1 at stage-1 and 0_2 at stage-2. We fix these as Nature's moves (mark them with the superscript S) in Γ given in Figure 2 to construct the $MCG(3)$ depicted here. All other information sets of Γ are decisive information sets of $MCG(3)$, and each of them is marked with D^3. To complete the LFE description, we need to solve for the LFE actions and beliefs at D^3, using the SE of $MCG(3)$. In the unique SE of $MCG(3)$, thus in LFE, at stage-3, all player-1 types choose $T3$ (asterisked), irrespective of beliefs. Thus 1_2 chooses $T2$ (asterisked) irrespective of beliefs. $2_1's$ stage-1 beliefs are derived from Nature's prior distribution using Bayes' rule. Thus, $2_1's$ stage-1 belief over $\{0_2, 1_2\}$ is $(\frac{1}{10}, \frac{9}{10})$ (asterisked). In SE, and thus in LFE, 2_1 chooses $T1$ (asterisked) because while $0_2's$ choice of $P2$ in stage-2 is fixed as Nature's move, 1_2 chooses $T2$ in stage-2. Thus $2_1's$ expected payoff from $P1$ is $\frac{16\times 1}{10} + \frac{2\times 9}{10} = 3.4$, which is strictly less than 4, his expected payoff from $T1$. Finally, consistency of SE beliefs implies $1_2's$ stage-2 LFE belief over $\{0_1, 1_1, 2_1\}$ is given by $(\frac{1}{2}, \frac{1}{2}, 0)$ (asterisked), and all player-1 types' stage-3 LFE belief over $\{0_2, 1_2\}$ is $(1, 0)$ (asterisked). To sum up, the LFE of this example (Figures 1-5) is as follows (π_{t_i} and μ_{t_i} denote $t_i's$ LFE strategy and LFE belief respectively).
$\pi_{0_1} = \pi_{1_1} = \{P1, T3\}$, $\pi_{2_1} = \{T1, T3\}$; $\pi_{0_2} = \{P2\}$, $\pi_{1_2} = \{T2\}$.
$\mu_{0_1} = \mu_{1_1} = \mu_{2_1} = \{(\frac{1}{10}, \frac{9}{10}), (1, 0)\}$ on $\{0_2, 1_2\}$ at stage-1 and at stage-3, respectively. $\mu_{0_2} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ and $\mu_{1_2} = (\frac{1}{2}, \frac{1}{2}, 0)$ on $\{0_1, 1_1, 2_1\}$ in stage-2.

on the behalf of his opponent in such a scenario. If he could apply greater foresight, he would also use it for himself! This has a parallel with Alaoui and Penta's (2016) model of Level-k. In particular, in their model if a player, say Bob, knows that his opponent, say Ann, is cognitively more sophisticated, and Bob knows that Ann knows this, then Bob chooses his own highest possible level. We now proceed to formally defining the Limited Foresight Equilibrium.

3.1 Limited Foresight Equilibrium: Definition

Consider an S-staged Interaction Game Γ. Γ generates S distinct Curtailed Games given the assumption that Nature's distribution over T is common knowledge in each Curtailed Game. Denote a strategy profile of Γ as π, where $\pi = ((\pi_{t_i})_{t_i \in N})$. Denote a belief system of Γ as μ, where $\mu = ((\mu_{t_i})_{t_i \in N})$. For each player-type t_i, (π_{t_i}, μ_{t_i}) specifies the action choice and belief of t_i at all the information sets of Γ where t_i moves. Formally, consider an arbitrary information set $I(t_i)$ where t_i moves. Let $I(t_i)$ be generated by the sequence $h_0 \in H_0$ from the underlying sequential game with perfect information. That is, $Seq^{-1}(I(t_i)) = h_0$ holds. Then $\pi_{t_i} : \; I(t_i) \longmapsto \Delta(A(I(t_i)))$, and $\mu_{t_i} : \; I(t_i) \longmapsto \Delta\{((t_i, t_{-i}), h_0) : t_{-i} \in T_{-i}\}$.

Consider a stage $n \in \{1, ...S\}$. Let (σ^n, b^n) denote an assessment for $CG(n)$. That is, $\sigma^n = ((\sigma^n_{t_i})_{t_i \in N})$ and $b^n = ((b^n_{t_i})_{t_i \in N})$ denote a strategy and belief profile for $CG(n)$ respectively. For each player-type t_i, $(\sigma^n_{t_i}, b^n_{t_i})$ specifies the (possibly mixed) action choice and beliefs of player-type t_i at all the information sets of $CG(n)$ where t_i moves. Let the set of Sequential Equilibria of any game G be denoted as $\Psi(G)$. Let D^n denote the decisive information sets of $CG(n)$. The Limited Foresight Equilibrium of Γ will be an assessment (π, μ) for Γ that we will construct below.

We need one more definition before defining an LFE: *Modified Curtailed Games* (*MCG* for short). A Modified Curtailed Game captures rule (a) in the rules-of-thumb. That is, a Modified Curtailed Game formally defines how to modify a Curtailed Game to model the feature that a limited-foresight type considers the LFE actions of lower types fixed as

Nature's moves.

Definition 4 *(Modified Curtailed Games)*: First, define $MCG(1) \equiv CG(1)$. Next, consider a Curtailed Game $CG(n)$ for some $n \in \{2, .., S\}$. Suppose $\pi(.)$ provides the LFE strategy profile for all the information sets in $\bigcup_{k=1}^{n-1} D^k$, which is the union of the decisive information sets of $CG(1)$ through $CG(n-1)$. Then $MCG(n)$ is defined by its construction from $CG(n)$, given π, by making *two modifications. First*, modify the player function of $CG(n)$, P^n, to mP^n so that in $MCG(n)$, for all the decisive information sets of $CG(1)$ through $CG(n-1)$, the player-type moving there is replaced by Nature. That is, in $MCG(n)$, we set $mP^n(I) = Nature$ for all the information sets $I \in \bigcup_{k=1}^{n-1} D^k$. For other information sets, the modified player function is the same as P^n, the player function of $CG(n)$. That is, for all $I \notin \bigcup_{k=1}^{n-1} D^k$, $mP^n(I) = P^n(I)$ holds. *Second,* we specify how Nature moves at these information sets using ρ^n, which is an augmented version of the Nature's move in $CG(n)$, given by ρ. In particular, in $MCG(n)$, the prior distribution over players' types is the same as $CG(n)$ and Γ, that is, $\rho^n(\emptyset) = \rho$. Additionally, for all $I \in \bigcup_{k=1}^{n-1} D^k$, $\rho^n(I) = \pi(I)$, that is, for all the decisive information sets of $CG(1)$ through $CG(n-1)$, Nature moves exactly as the LFE action specified by π.

In what follows we will assume the following. Assumption 1 (below) imposes a restriction on the prior distributions over foresight-levels across players.

Assumption 1: The common-knowledge prior distribution over the foresight-levels across players, ρ, is such that for every player i and for every i-type, t_i, conditional on i's type being t_i, the probability of the profile $(S - s_j)_{j \neq i}$[23] is strictly greater than 0.

$$\Pr((t_j \text{ has no foresight limitation})_{j \neq i} | \rho, t_i) = \Pr((t_j = (S - s_j)_j)_{j \neq i} | \rho, t_i) > 0,$$

[23] $(S - s_j)_j$ is the player-j-type with no foresight limitation.

$$\forall i \in N_0,\ \forall t_i \in T_i.$$

Assumption 1 states that we only allow those prior distributions over foresight-levels which imply that every player-type has a strictly positive prior probability of facing an opponent-profile such that each of the opponents is of the type who has no foresight limitation. We will discuss the point at which Assumption 1 is used after defining the LFE.

Given the notation and definitions above, we have the following definition of LFE.

Definition 5: *(π, μ) is a **Limited Foresight Equilibrium** of an S-staged Interaction Game Γ if it is constructed in the following S steps:*

Step 1: Select a Sequential Equilibrium assessment $(\sigma^1, b^1) \in \Psi(CG(1))$. Set $(\pi(I), \mu(I)) = (\sigma^1(I), b^1(I))\ \forall I \in D^1$.

Step 2: Convert $CG(2)$ to $MCG(2)$ using $\pi(D^1)$ obtained from Step 1. Select an assessment $(\sigma^2, b^2) \in \Psi(MCG(2))$. Set $(\pi(I), \mu(I)) = (\sigma^2(I), b^2(I))\ \forall I \in D^2$.

Step n: Convert $CG(n)$ to $MCG(n)$ using $\pi(\bigcup_{k=1}^{n-1} D^k)$ obtained from Step 1 through Step (n-1). Select an assessment $(\sigma^n, b^n) \in \Psi(MCG(n))$. Set $(\pi(I), \mu(I)) = (\sigma^n(I), b^n(I))$ $\forall I \in D^n$. Repeat Step n until $n = S$.[24]

4 Limited Foresight Equilibrium Properties

Remark 1: The Interaction Game nests the underlying finite sequential game with perfect information. In particular, if an Interaction Game, Γ, has the common knowledge prior

[24] In the construction of an LFE in Definition 5, we are assuming that a player-type observing a longer CG correctly anticipates which one of the many possible SE was selected at each of the shorter Curtailed Games. For example, for constructing $MCG(2)$, we need a selection from the Sequential Equilibria of $CG(1)$. We are assuming that 0_2 at stage 2 and 1_1 at stage 1 correctly guess which one of the many possible optimal choices is chosen by 0_1 at stage 1. This assumption is significant in general and comes with the territory of an equilibrium notion. This assumption has no bearing on our Centipede game and Sequential Bargaining game results, as any $CG \neq \Gamma$ has a unique SE there.

distribution, ρ, such that all players are of the no-foresight-limitation type with a prior probability equal to 1, that is, for all $i \in N_0$, $Pr(t_i|\rho) = 1$ holds iff $t_i = (S - s_i)_i$ holds, then Γ is equivalent to Γ_0, the underlying sequential game with perfect information that generated Γ.[25] Therefore, in that case, the set of LFE of Γ is equal to the set of Sequential Equilibria of Γ which is identical to the set of SPNE of Γ_0.

Now we describe some properties of the LFE defined above. First, existence and upper hemi-continuity.

Proposition 1(a) *(existence)*: *For every finite Interaction Game, there exists at least one Limited Foresight Equilibrium.*

1(b) *(upper hemi-continuity): Given the extensive form,* $\{N, H, \{\mathcal{I}(t_i)\}_{t_i \in N},\ P, A\}$ *, for an Interaction Game, the correspondence from pairs* (ρ, u) *of initial probability distributions and payoff profiles to the set of Limited Foresight Equilibria for the Interaction Game so defined is upper hemi-continuous.*

The proof of existence and upper hemi-continuity of LFE follows from the existence and upper hemi-continuity of the Sequential Equilibrium (Kreps and Wilson (1982b)). The details are given in the Appendix. Note 1 (below) characterizes trivial information sets where the LFE, (π, μ), places no restrictions on beliefs or actions.

Note 1: Trivial information sets. Fix an LFE, (π, μ), of an S-staged Interaction Game Γ. Suppose t_i is moving at a stage-K information set: $I(t_i) = \{((t_i, t_{-i}), (a_0^k)_{k=1,\dots,K-1})\colon\ t_{-i} \in T_{-i}\}$. Suppose for some $r < K$, there exists a subsequence $(a_0^k)_{k=0,\dots,r-1}$ of $(a_0^k)_{k=1,\dots,K-1}$, such that at the information set $I'(t_i) = \{((t_i, t_{-i}), (a_0^k)_{k=0,\dots,r-1})\colon\ t_{-i} \in T_{-i}\}$, t_i moves and the total foresight of t_i at $I'(t_i)$ is strictly less than S.[26] Given these conditions, if $\pi_{t_i}(a_0^r|I'(t_i)) = 0$

[25] For example, Palacios-Huerta and Volij (2009) may have been able to establish this condition in their experiment when expert chess players played other expert chess players in a Centipede game, and the degree of expertise of both players in each pair was common knowledge among the pair.

[26] Here and in what follows, (t, a_0^0) is defined as the stage-1 move after Nature chooses the player-type combination $t \in T$.

holds, then $I(t_i)$ occurs with probability 0 because t_i's own prior move, at an information set where his total foresight was strictly less than S, rules out the possibility of reaching $I(t_i)$. Thus, $I(t_i)$ is a trivial information set of $MCG(K+t_i)$,.., $MCG(S)$ and the properties of LFE are irrelevant there.[27]

Note 1 states that when we construct the $MCG(K+t_i)$, t_i's move at the preceding information set will be considered as Nature's move, and if Nature's move rules out $I(t_i)$, then it is theoretically and empirically a zero probability event. $I(t_i)$ being a trivial information set implies that if we modify the LFE (π, μ) by arbitrarily changing the action or beliefs only at $I(t_i)$ to generate (π', μ'), then (π', μ') is also an LFE. As an example, suppose in the Centipede game example in Figures 1-5, the payoff profile from $T1$ is changed to $(10, 1)$. Then 0_1's LFE action at stage-1 will be $T1$. So, 0_1's stage-3 information set will be a trivial information set, because his own action of $T1$ rules out the action sequence $(P1, P2)$ being played. Note that, given Assumption 1, the only way trivial information sets of a player-type are generated is by the player-type's own LFE action, at a preceding information set, where that player-type's total foresight is strictly less than S. This is because, by Assumption 1, in every MCG, the profile of opponent-types with no ex-ante foresight limitation, $((S-s_j)_j)_{j \neq i}$, occurs with a strictly positive prior probability, and none of their moves are fixed as Nature's moves in any MCG. So, in what follows, the properties of LFE described below are relevant only at non-trivial information sets.

Next, we define the *corresponding information sets* of different types of the same player as follows. For the purpose of this definition, consider a player $i \in N_0$ and any two i-types, t_i and t_i', such that $t_i \neq t_i'$. Consider two information sets of an Interaction Game, $I(t_i)$ and $I(t_i')$, such that the player-type moving at $I(t_i)$ is t_i, and the player-type moving at $I(t_i')$ is t_i'.

Definition 6 *(corresponding information sets)*: Any two information sets, $I(t_i)$ and $I(t_i')$

[27]Here and in what follows, if $(K+t_i) > S$, then $MCG(K+t_i) \equiv MCG(S)$.

are *corresponding information sets* if they are preceded by the same action sequence, except Nature's initial move at the $\emptyset$ history. That is, if $Seq^{-1}(I(t_i)) = Seq^{-1}(I(t_i')) = h_0$, for some action sequence $h_0 \in H_0$ of the underlying game, then $I(t_i)$ and $I(t_i')$ are *corresponding information sets*. In other words, for some $h_0 \in H_0$, $I(t_i)$ and $I(t_i')$ are corresponding information sets if $I(t_i)$ is of the form $I(t_i) = \{((t_i, t_{-i}), h_0) : \ t_{-i} \in T_{-i}\}$ and $I(t_i')$ is of the form $I(t_i') = \{((t_i', t_{-i}), h_0) : \ t_{-i} \in T_{-i}\}$.

At corresponding information sets, different types of the same player move after observing the same sequence of actions. Using this definition, we will now state Proposition 2 and its corollary. Suppose the prior probability distribution over the foresight-levels, ρ, is pairwise-independent across players. Then Proposition 2 and its corollary tell us that calculating the SE of a given $MCG(n)$ is typically simple. In particular, Proposition 2 implies that if ρ obeys the pairwise-independence property, then for any SE of any $MCG(n)$, the SE beliefs (not necessarily LFE beliefs) of all types of a particular player at corresponding (non-trivial) information sets are identical. That is, if different types of a particular player observe the same prior sequence of actions, then, in SE, they have identical beliefs over the possible profiles of opponents. The corollary of this proposition is that if an SE of $MCG(n)$ specifies a strategy for a player-type that is strictly better than alternative strategies from a particular information set of $MCG(n)$, then the SE must specify an identical strategy for other types of that player from their respective corresponding information sets. We now formally state Proposition 2 and its corollary.

Proposition 2 *(SE are "simple"): Consider an arbitrary $MCG(n)$ with a pairwise independent prior distribution over players' types. That is, $Pr(t_i, t_j|\rho) = Pr(t_i|\rho).Prob(t_j|\rho)$ for any $t_i \in T_i$ and $t_j \in T_j$ and for any $i, j \in N_0$. Let σ^n be some totally mixed strategy profile of $MCG(n)$. Let $I(t_i)$ and $I(t_i')$ be (non-trivial) corresponding information sets which are preceded by the same action sequence (excluding Nature's initial move at $\emptyset$), given by h_0. If* *beliefs $b^n(.)$ are calculated using σ^n and Bayes' rule, then:*

$$b^n((t_{-i}, h_0)|\ t_i,\ I(t_i),\ \sigma^n) = b^n((t_{-i}, h_0)|\ t_i',\ I(t_i'),\ \sigma^n)\ \textit{holds}\ \forall t_{-i} \in T_{-i}. \tag{2}$$

Therefore, for any MCG(n) and for any of its Sequential Equilibria, for all $i \in N_0$, the equilibrium beliefs of all types of player i are identical at corresponding information sets.

The proof is given in the Appendix.

Corollary of Proposition 2: Consider an arbitrary $MCG(n)$ and a SE, (σ^n, b^n), of $MCG(n)$. Suppose ρ satisfies the independence property described in Proposition 2. Consider an arbitrary action sequence h_0, which immediately follows Nature's initial move at the $\emptyset$ history. Suppose that σ^n specifies a strict best response for t_i from each information set following h_0. Then σ^n must specify the same strategy for all t_i' at each corresponding information set of t_i' following the action sequence h_0.

The corollary of Proposition 2 follows because, by Proposition 2, other types of player i have identical beliefs at corresponding information sets. Further, they face the same strategy profile $(\sigma^n_{t_j})_{j\neq i}$. Therefore, at corresponding information sets, the mapping from strategies to expected payoff is the same for all $i - types$. As σ^n specifies a strict best response for t_i from each information set following h_0, the sequential rationality requirement of a SE implies $\sigma^n_{t_i'} = \sigma^n_{t_i}$, for all $t_i' \in T_i$, at corresponding information sets of t_i' following h_0. The proof is in the Appendix.

Illustrating Proposition 2. Consider $CG(1) \equiv MCG(1)$ in Figure 3; note that the SE beliefs of all player-1 types at corresponding information sets in stage-one are the same: $(\frac{1}{10}, \frac{9}{10})$ on $\{0_2, 1_2\}$. Consider $MCG(2)$ in Figure 4, note that the SE beliefs of 0_2 and 1_2 at corresponding information sets in stage-two are the same; $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ on $\{0_1, 1_1, 2_1\}$. Further, in $MCG(2)$, the SE beliefs of 1_1 and 2_1 at corresponding information sets in stage-one are also the same: $(\frac{1}{10}, \frac{9}{10})$ on $\{0_2, 1_2\}$. Last, consider $MCG(3)$ in Figure 5; note that the SE beliefs of all player-1 types at corresponding information sets in stage-three are the same: $(1, 0)$ on $\{(\text{Nature as } 0_2), 1_2\}$. It is worth noting that the corollary of Proposition 2 also holds in each case. In the SE of $CG(1)$, all 1-types choose $P1$ at corresponding information sets in stage-one. In the SE of $MCG(2)$, both 2-types choose $P2$ at corresponding information sets in stage-two and both 1_1 and 2_1 choose $P1$ at corresponding stage-one information sets. Last, in $MCG(3)$, all 1-types choose $T1$ at corresponding information sets in stage-three.

Discussion of Proposition 2. Recall that the purpose of constructing an $MCG(n)$ is to model how a player-type, say t_{Ann}, moving at a decisive information set of $MCG(n)$, perceives the Interaction Game and optimizes. In $MCG(n)$, all player-types with total foresight less than n have been replaced by Nature and their moves are commonly known as Nature's moves. To optimize, t_{Ann} has to anticipate what other player-types will choose in $MCG(n)$. The SE selected for this $MCG(n)$ represents what t_{Ann} thinks (and all player-types at the other decisive information sets of $MCG(n)$ think) will be the strategies and beliefs of all player-types in $MCG(n)$. Using the SE implies that t_{Ann} thinks that all player-types will play sequentially rational strategies in $MCG(n)$ and that they will have beliefs consistent with the strategy profile so constructed. In principle, the SE of $MCG(n)$ can be quite complicated. However, Proposition 2 states that, when the prior distribution over types is independent across players, the SE of $MCG(n)$ boils down to t_{Ann} "lumping" different higher types of the same player into one. That is, in $MCG(n)$, if two different types of a player observe the same prior action sequence, then in any SE (t_{Ann} thinks that) they have the same belief over opponents' type profile. This identical belief property has the corollary that if in SE, a player-type's chosen strategy, $\sigma^n_{t_i}$, from a given information set is strictly better than the next-best alternative, then in the SE of $MCG(n)$, (t_{Ann} thinks that) the strategy of all other types of that player at their corresponding information sets is also $\sigma^n_{t_i}$. That is, (t_{Ann}

thinks that) *all higher types of a particular player can be treated identically in $MCG(n)$ up to the case of indifference.*[28] Next, starting with Note 2, we move on to specifying other properties of LFE beliefs.

Note 2: *(LFE beliefs of player-types with no foresight limitation).* Suppose that t_i has no foresight limitation at $I(t_i)$. That is, suppose the total foresight of t_i at $I(t_i)$ is S (the number of stages in the Interaction Game, Γ). Then the LFE beliefs of t_i at $I(t_i)$ are calculated using some SE, (σ^S, b^S) of $MCG(S)$. In $MCG(S)$, the LFE strategy profile, π, at $\bigcup_{n=1}^{S-1} D^n$ are given as Nature's moves. Further, π at D^S is given by σ^S at D^S. So, Bayes' rule implies,

$$\mu_{t_i}(h|\ I(t_i)) = \frac{Pr(h \mid \pi, \rho)}{Pr(I(t_i) \mid \pi, \rho)} \ \textit{holds}\ \forall h \in I(t_i),\ \textit{if}\ Pr(I(t_i) \mid \pi, \rho) > 0.$$

If $Pr(I(t_i) \mid \pi, \rho) = 0$, then $\mu_{t_i}(h|\ I(t_i))$ is equal to the SE beliefs $b^S_{t_i}(h|\ I(t_i))$, $\forall h \in I(t_i)$.[29]

Note 2 tells us that the LFE beliefs of player-types with no foresight limitations are standard. We now describe properties of LFE beliefs of player-types at moves where they *may have* foresight limitations. Consider an arbitrary player-type, t_i, moving at an information set $I(t_i)$. Recall that $I(t_i)$ is identified by the player-type moving there, t_i in this case, and the action sequence of the underlying sequential game with perfect information that generates it, say $h_0 = (a^k_0)_{k=1,\dots,K-1}$ such that $Seq^{-1}(I(t_i)) = h_0$ holds. In what follows we will refer to histories/sequences in information sets as "nodes." So if $I(t_i) = \{((t_i, t_{-i}), h_0)\colon t_{-i} \in T_{-i}\}$, then each different node of $I(t_i)$ implies a different combination of opponents' types, $t_{-i} \in T_{-i}$, that, along with t_i, played the same preceding action sequence: h_0. In what

[28] The case of indifference doesn't arise in any MCG shorter than the Interaction Game in the Bargaining game and the Centipede game applications. Note that we have not been able to show the following claim: For any MCG, there always exists a SE such that different types of a player have identical strategies and beliefs at corresponding information sets. The notation of the possible proof of this claim is turning out to be too complicated for this paper. However, Proposition 2 already shows this property for SE beliefs. Further, the proof of the corollary of Proposition 2 shows the equality of the expected payoff of different types of a player from identical strategies at corresponding information sets. Thus, the marginal gain from the claim appears to be limited.

[29] The definition of a SE implies that $b^S_{t_i}(h|\ I(t_i))$ is well defined $\forall h \in I(t_i)$ when $Pr(I(t_i) \mid \sigma^S, \rho^S) = Pr(I(t_i) \mid \pi, \rho) = 0$.

follows, it will be useful to distinguish among nodes of $I(t_i)$ on the basis of opponents'-types' total foresight. In particular, it will be useful to identify those nodes of $I(t_i)$ such that at each move in the action sequence preceding that node, the opponent-type moving there had strictly lesser total foresight (was a lower type). We will call the subset of $I(t_i)$ containing such nodes as $L(I(t_i))$. We now define this set.

Definition 7 (*playing against lower types*): Consider an arbitrary stage-K information set $I(t_i)$ of Γ. Let $I(t_i) = \{(t,\ (a_0^k)_{k=1,\ldots,K-1}) : t_{-i} \in T_{-i}\}$; that is, let $(a_0^k)_{k=1,\ldots,K-1}$ be the action sequence preceding $I(t_i)$ (excluding Nature's initial move at the $\emptyset$ history). Consider h, an arbitrary node of $I(t_i)$ of the form $h = (t,\ (a_0^k)_{k=1,\ldots,K-1})$. Define $L(I(t_i))$ as the subset of $I(t_i)$ such that a node h in $I(t_i)$ belongs to $L(I(t_i))$ iff for any subsequence $\hat{h} = (t,\ (a_0^k)_{k=0,\ldots,r-1})$ of h, the opponent-type moving at $\hat{h}$ has strictly lower total foresight at $\hat{h}$ than t_i's total foresight at h. That is, if $P(\hat{h}) = t_j$, then the total foresight of t_j at $\hat{h}$, given by $t_j + r$, is strictly less than $(t_i + K)$, which is t_i's total foresight at h.[30]

Definition 7 states that for each $h \in L(I(t_i))$, t_i, is playing against opponents'-types who were lower types at all preceding moves. In other words, at h, t_i has strictly greater total foresight than the total foresight of his opponents at their respective prior moves along the action-sequence $(a_0^k)_{k=1,\ldots,K-1}$ leading into $I(t_i)$. Define the complement of $L(I(t_i))$ as $L^c(I(t_i)) = I(t_i) - L(I(t_i))$. At any node in $L^c(I(t_i))$, t_i has weakly lower total foresight than at least one of his opponents at some move of that opponent along the action-sequence $(a_0^k)_{k=1,\ldots,K-1}$. For the two-player case, a node in $L^c(I(t_i))$ implies that the opponent is a higher type at some move preceding that node.

Proposition 3 (below) states a consistency condition that the LFE belief, μ, obeys. It states that in any LFE $(\pi,\ \mu)$, given an information set, I, the belief distribution over $L(I)$, which contains the nodes of I where the player-type moving at I is playing against lower

[30] If $h \in L(I(t_i))$ and $(t_i + K) \geq S$ hold, then $(t_j + r)$ must be strictly less than S.

opponents'-types, should be derived from the LFE strategy profile, π, and the prior distribution, ρ, using Bayes' rule wherever possible. In other words, the LFE belief distribution over $L(I)$, conditional on $L(I)$, is "accurate" with respect to the LFE strategy profile π. Remark 2 leads us to an important implication of Proposition 3.

Proposition 3: *Let I be a stage-K information set of Γ, such that $P(I) = t_i$. In any LFE (π, μ), t_i's belief distribution over the set of nodes in $L(I)$, conditional on $L(I)$, must be derived from the LFE strategy profile using Bayes' rule wherever possible. That is, $\forall h \in L(I)$, if $Prob(L(I) \mid \rho^{K+t_i}) > 0$ holds, then we must have:*

$$\mu_{t_i}(h|\ L(I)) = \frac{Pr(h \mid \pi, \rho)}{Pr(L(I) \mid \pi, \rho)} = \frac{Pr(h \mid \rho^{K+t_i})}{Pr(L(I) \mid \rho^{K+t_i})}\ \forall h \in L(I). \tag{3}$$

Remark 2: If player i's foresight-level is higher, say t_i' instead of t_i, then for any action sequence observed, more opponent-type combinations, $t_{-i} \in T_{-i}$, are such that along the sequence of actions observed, i's opponents'-types playing the actions are lower types (whose LFE strategies at preceding moves are known to i's type) compared to i's type. That is, consider two corresponding information sets, $I(t_i)$ and $I(t_i')$, of Γ. If $t_i < t_i'$ holds, then we must have $|L(I(t_i))| \leq |L(I(t_i'))|$.

The proofs of Proposition 3 and Remark 2 are given in the Appendix. Proposition 3 follows from the recursive construction of an LFE. As an example of Proposition 3, suppose *Ann* and *Bob* are playing a 6-staged alternate move game where *Ann* moves first. Suppose 1_{Ann} observes a sequence of two actions at her stage-3 information set, I. Suppose according to the LFE strategy profile and ρ (the prior distribution over $Ann's$ types and $Bob's$ types), conditional on I, 1_{Ann} faces 0_{Bob}, 1_{Bob} and 2_{Bob} with probabilities $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ respectively. Note that 0_{Bob} and 1_{Bob} at stage-2 are lower types than 1_{Ann} at stage-3. Proposition 3 states that the LFE belief of 1_{Ann} must satisfy:

$$\mu_{1_{Ann}}(\text{Bob's type is } 0_{Bob}|\text{ Bob's type is } 0_{Bob} \text{ or } 1_{Bob},\ I) = \frac{1}{3}, \text{ and}$$

$$\mu_{1_{Ann}}(\text{Bob's type is } 1_{Bob}|\text{ Bob's type is } 0_{Bob} \text{ or } 1_{Bob},\ I) = \frac{2}{3}.$$

However, Proposition 4 (stated below) will tell us that $\mu_{1_{Ann}}(\text{ Bob's type is } 0_{Bob} \text{ or } 1_{Bob}|\ I)$ need not by equal to $\frac{3}{4}$, as implied by the LFE strategy profile and ρ. This is because in LFE, 1_{Ann} is allowed to "misunderstand" higher *Bob-types'* LFE strategies. Given Proposition 3, the significance of Remark 2 is that, given a sequence of prior moves observed, if a player's foresight-level is higher, then the player's beliefs are "accurate," in the sense of Proposition 3, over more opponent-type combinations. In the present example, Proposition 3 implies that at the corresponding information set of 2_{Ann}, say $I(2_{Ann})$, the LFE belief of 2_{Ann}, $\mu_{2_{Ann}}(m_{Bob}|\ I(2_{Ann}))$, must be accurate with respect to the LFE strategy profile and ρ for $m = 0$, 1, and 2.

Remark 2 approximately captures the findings from Reynolds (1992). Reynolds (1992), while testing recognition of opponent's expertise among chess players, found that "Higher rated players consistently made lower estimation errors" (of other chess players' ELO ratings). If one proxies for foresight using experience-level or ELO ratings, then the Remark 2 hints at this finding. The reasons for only approximate similarity to Reynolds' (1992) finding are: First, the proxying of foresight using ELO ratings is a leap of faith. Second, in an LFE, the *total* belief placed on lower types, $\mu_{t_i}(L(I(t_i))|\ I(t_i))$, need not be derived from the LFE strategy profile using Bayes' rule (shown in Proposition 4 (below)). However, as Proposition 3 says, conditional on $L(I(t_i))$, the distribution of $\mu_{t_i}(L(I(t_i))|\ I(t_i))$ among the various nodes of $L(I(t_i))$, i.e. the distribution of $\mu_{t_i}(L(I(t_i))|\ I(t_i))$ among lower types, is derived from the LFE strategy profile using Bayes' rule. It is notable that starting from the same common knowledge belief over opponents' types, the belief of higher foresight-level types becomes "more accurate" (at least in the sense of Proposition 3 and Remark 2) after the same sequence of actions. Proposition 4 specifies the limited accuracy of LFE beliefs.

Proposition 4: *Fix an LFE, (π, μ), of Γ. If the total foresight of t_i at $I(t_i)$ is strictly less than the number of stages in the Interaction Game, then the beliefs of t_i over the nodes in $I(t_i)$ need not be derived from the LFE strategy profile using Bayes' rule. Thus,* ***it need not be true that*** $\mu_{t_i}(h|\ I(t_i)) = \frac{Prob(h|\ \pi,\rho)}{Prob(I(t_i)|\ \pi,\rho)}$ *holds* $\forall\, h \in I(t_i)$ *and for all (non-trivial) information sets* $I(t_i)$ *of* Γ .

Proof of Proposition 4 by counter-example. Consider the Interaction Game in Figure 2. The LFE for this Interaction Game is specified in the notes of Figure 5. The second-stage LFE belief of 0_2 after observing $P1$ in the first-stage is $\mu_{0_2} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ over $\{0_1, 1_1, 2_1\}$. However, 2_1, in LFE, puts 0 probability on $P1$. So according to the LFE strategy profile and Bayes' rule, after observing $P1$, the belief of 0_2 over $\{0_1, 1_1, 2_1\}$ should be $(\frac{1}{2}, \frac{1}{2}, 0) \neq (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. *Q.E.D.*

Proposition 4 follows because the LFE beliefs at a given information set are derived using a SE of that information set's decisive Modified Curtailed Game. The selected SE strategy profile for this MCG may not stipulate the same actions as the LFE strategy profile at the non-decisive information sets of the MCG. Recall that, although the SE of the MCG and the LFE strategy profiles can be different, the SE strategy profile provides a sequentially rational best response strategy in the MCG for each player-type given the SE beliefs.

Proposition 5 (below) states that, in LFE, if a player-type observes a sequence of moves that cannot occur when playing against lower opponents'-types (whose moves are considered fixed as Nature's moves), then he discovers that he is playing against at-least one higher opponent-type, and must use his total foresight at that move to optimize.

Proposition 5: *Suppose Assumption 1 holds. Consider a stage-K information set,* $I(t_i)$*, of* $MCG(K+t_i)$*. Suppose Nature's moves in* $MCG(K+t_i)$*, denoted by* ρ^{K+t_i}*, imply that there is 0 probability of reaching the nodes in* $L(I(t_i))$*, which have the property that Nature moves* *at all moves in the sequences preceding these nodes.*[31] *Then for any LFE,* (π, μ)*, the LFE belief of* t_i*, conditional on* $I(t_i)$*, must put probability 1 on the nodes in* $L^c(I(t_i))$*, which have the property that at some move in the preceding sequence, some opponent-type moving there had total foresight of at least* $min\{(K + t_i), S\}$*. That is, for all* $t_i \in N$*, for all* $I(t_i) \in \mathcal{I}(t_i)$ *(such that* $I(t_i)$ *is a non-trivial information set):*

$$[Prob(L(I(t_i)) \mid \rho^{K+t_i}) = 0] \implies \mu_{t_i}([L(I(t_i))]^c \mid I(t_i)) = \frac{Prob([L(I(t_i))]^c)| \ \rho, \pi)}{Prob(I(t_i)| \ \rho, \pi)} = 1 \quad (4)$$

The proof is given in the Appendix.[32] Consider t_i at an information set $I(t_i)$ which is generated from some action sequence h_0 of the underlying game. Proposition 5 takes Proposition 3 further. According to Proposition 3, t_i knows the probability with which lower opponents'-types choose the actions in the action sequence h_0. In fact, t_i knows these lower opponents'-types' moves as Nature's moves. So if t_i knows that the choices of these lower opponents'-types imply a zero probability of h_0 being played and yet he finds himself moving after h_0, then he knows that he is at a node of $I(t_i)$ where at least one opponent-type was a higher type at some preceding move. In the two player case it means that one knows that one's opponent had a higher total foresight at some preceding move. Recognition of the higher opponent-type implies that the LFE actions of the lower types don't matter for t_i's calculation of the sequentially rational action at $I(t_i)$; t_i must use his total foresight to optimize.[33]

[31] Here and in what follows, if $(K + t_i) > S$, then $\rho^{(K+t_i)} \equiv \rho^S$.

[32] We use Assumption 1 to prove Proposition 5. Assumption 1 ensures that for all limited-foresight player-types, higher opponent-types exist with a strictly positive probability.

[33] As an example of Proposition 5, suppose that in the 3-staged Centipede Game (Figure 1), we changed the payoff profile from $T1$ to (10,1). Then in the corresponding $MCG(2)$, when 0_2 observes $P1$, in SE (and thus in LFE) 0_2 must put probability 1 on his opponent being 1_1 or 2_1, because 0_2 has fixed 0_1's choice of "$T1$ with probability 1" as Nature's move in $MCG(2)$. Note that Assumption 1 is important here. If 0_2 had put 0 prior probability over both 1_1 and 2_1, then the LFE belief of 0_2 after observing $P1$ would not be well defined. Assumption 1 guarantees that 0_2 puts a strictly positive prior probability on his opponent being 2_1 and thus, his SE (and thus LFE) beliefs are well defined.

Reputation Concerns. It is worth noting that using the SE to solve for the LFE actions at decisive information sets of appropriate Modified Curtailed Games implies that player-types take reputation concerns into account in order to optimize within the MCG (which may be shorter or equal to the Interaction Game). As the propositions above specify, in LFE, player-types are updating about the opponents' foresight-levels within the play of the game. So player-types must choose the optimal between mimicking a lower type or choosing a strategy that is different from the lower types which may reveal his/her own foresight-level to opponents. For example, in the next section, we will see that in the Centipede game, even player-types with no foresight limitation pretend to be a limited-foresight type to gain. One can think of several applications where it is better to reveal one's high foresight-level. For example, suppose a firm is looking for a partner for starting a new business. Further, suppose the new business requires several stages of costly and unrecoverable investment from both partners before "high" returns accrue to both parties. The firm will look for a partner who has displayed high foresight in previous interactions, otherwise it runs the risk of a low-foresight partner quitting midway and saddling it with losses.[34] Thus it would be optimal to reveal one's high foresight-level to the firm if one is sufficiently confident of the firm's foresight. Remark 3 (below) specifies the effect on the LFE beliefs of a player-type from observing more/less preceding actions.

Remark 3: Suppose t_i moves at two information sets at stage K and K' of Γ. Suppose $K < K'$ holds. Then, in LFE, by the construction of LFE, if t_i at stage-K knows (as Nature's moves) the LFE action choice of some player-type, say t_j, at $I(t_j)$, then t_i, moving at stage-K' also knows the LFE action choice of t_j at $I(t_j)$; but the converse may not hold.

Remark 3 follows because at stage-K', t_i knows players'-types' LFE strategy at the information sets $\bigcup_{n=1}^{(K'+t_i-1)} D^n$. At stage-$K$, t_i knows players'-types' LFE strategy at $\bigcup_{n=1}^{(K+t_i-1)} D^n$; and $\bigcup_{n=1}^{(K+t_i-1)} D^n \subset \bigcup_{n=1}^{(K'+t_i-1)} D^n$ holds. Remark 3 approximately mirrors findings from

[34] It is easy to construct payoffs where the $\frac{min+max}{2}$ curtailment rule implies positive payoffs in the first few stages of such a game, but negative curtailed payoffs after these stages.

Reynolds (1992) and Rampal (2017). The latter study found that the more moves of the opponent observed by an expert "race game" player, the better his/her guess about the opponent's experience-level. In the same token, Reynolds (1992) found that estimation error about the opponent's ELO rating (as a chess player) decreased as a function of number of moves revealed. Remark 3 suggests that in an LFE this can happen, because when moving at a later stage, the same player-type has higher total foresight and hence observes a longer Modified Curtailed Game. Thus, the player-type knows (as Nature's moves) other player-types' moves he knew at the earlier stage, and, in most cases he knows more moves of other player-types at the later stage.

5 Applications

In this section we apply the LFE model to the Centipede game introduced by Rosenthal (1981) and the Sequential Bargaining game analyzed by Rubinstein (1982) and Ståhl (1972). We do not add game-specific types (eg. altruistic types, or other "crazy" types) to explain the qualitative findings for these games. Such additional features are often important in better capturing stylized data facts of different games. Instead, we restrict ourselves to the LFE model where players are selfish utility maximizers, but can possess various degrees of limited-foresight; and limited-foresight player-types are uncertain about his/her opponents' foresight-levels. The aim of this section is to illustrate general applicability of this LFE model; and to show that even without additional features, the LFE model explains various qualitative findings on two of the more studied empirical puzzles in perfect information sequential games. To see an experimental evaluation of the LFE model vis-a-vis other behavioral models, we direct the interested reader to Rampal (2017) which uses "race games" to illustrate the novel predictions of LFE.

5.1 Sequential Bargaining

The Sequential Bargaining game (Rubinstein (1982) and Ståhl (1972)) has been studied extensively in the literature (c.f. Binmore et al (1985), Neelin et al (1988), Guth and Tietz (1988,1990), Ochs and Roth (1989), Johnson et al (2002), and Binmore et al (2002), among others). One version of the game consists of two players bargaining over a pie of size X over multiple periods. In each period one player makes a proposal on how to split the pie, and the other player accepts or rejects this proposal. So each period has two stages: Proposal stage followed by accept/reject decision stage. If a proposal is accepted then the game ends and that proposal is implemented. If a proposal is rejected then the game proceeds to the next period where the player who rejected the last proposal now makes an offer, but from a smaller pie as the pie gets multiplied by a "common" discount factor, $\delta \in [0, 1]$. In the finite period case, if no proposal is accepted, then after a rejection in the last period, both players get 0 payoff. The SPNE prediction is that in a K period bargaining game, when K is odd, the first proposal which offers the first-mover/proposer $X[(1-\delta)\frac{(1-\delta^{K-1})}{1-\delta^2} + \delta^{K-1}]$ will be accepted.

Four stylized data trends, which are incongruent to the SPNE outcome prediction, have emerged in the experimental study of the Sequential Bargaining game. First, a tendency for first offers proposing equal split (Guth and Tietz (1988); Ochs and Roth (1989)) or offering the second round pie (Neelin et al (1988)) to the second-mover. Second, offers made in the first period are often rejected (Ochs and Roth (1989)). Third, and perhaps the most surprising finding is that the first period offers are very often succeeded by **disadvantageous counteroffers** (Ochs and Roth (1989) found that 81 percent of counteroffers were disadvantageous). Fourth, subgame consistency is violated (Binmore et al (2002)): the outcomes of an s-period Bargaining game are different from the outcomes of a "theoretically identical" s-period Bargaining *subgame* of an S-period Bargaining game (where $S > s$).

Now, we proceed to show that one can rationalize all these stylized data facts simultaneously by utilizing the LFE model of limited foresight and uncertainty about the opponent's

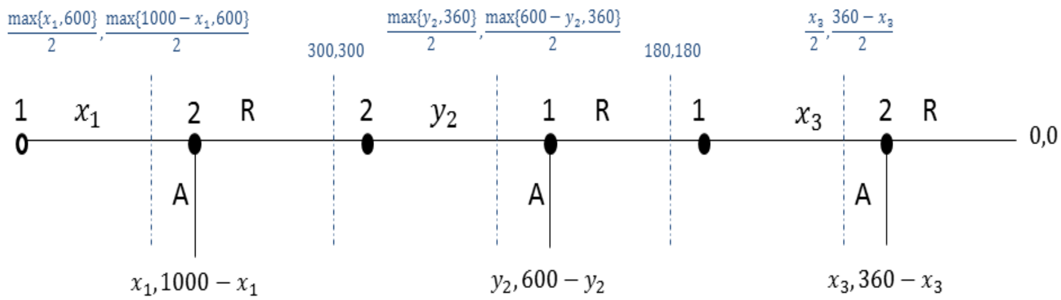

Figure 6. Sequential Bargaining Game and Associated Curtailed Payoffs Without Uncertainty

Notes: The figure shows curtailed payoff profiles being calculated using the (min+max)/2 curtailment rule. The curtailed payoffs are depicted in blue above the game. The pies are 1000, 600 and 360 in period 1, 2 and 3 respectively. x_i is player 1's offer (to himself) in period i. y_2 is player 2's offer (to player 1) in period 2. R implies "reject" and A implies "accept."

foresight. In particular, our rationalization does not use altruistic preferences, or preferences specific to the Bargaining game.

We consider the three-period Bargaining game with $\delta = 0.6$ as Γ_0, the underlying sequential game with perfect information.[35] We make the initial size of pie 1000 for simplicity. Note that the SPNE outcome is for the first mover to propose a split of $(760, 240)$ in the first period and for the second mover to accept it immediately. We convert Γ_0 into the Interaction Game, Γ, and present the features of its LFE. Figure 6 depicts the Curtailed payoffs associated with Γ_0, without showing the uncertainty.

As the three-period Bargaining game given in the Figure 6 has six stages, we have six player-1 types $\{0_1,\ 1_1,\ 2_1,\ 3_1,\ 4_1,$ and $5_1\}$ and five player-2 types $\{0_2,\ 1_2,\ 2_2,\ 3_2,\ 4_2\}$. For simplicity, we assume independent uniform prior distributions on both players' types. The LFE strategies for this uniform case are calculated in the Appendix and detailed in Table 1 there.[36]

[35]These specifications are used by Ochs and Roth (1989) in one of their treatments. Neelin et al (1988) and Johnson et al (2002) use $\delta = 0.5$, which doesn't change the features of the LFE outcome we discuss below.

[36]In practice, the prior distribution, ρ, is a free parameter available to the researcher that can be optimized to maximize the likelihood of the observed data. Rampal (2017) does this for the observed data on "race games."

The following qualitative outcomes observed by the studies on the Bargaining game (mentioned in brackets) are observed in the LFE that we detail in the Appendix. We now state brief explanations of how and why the LFE generates these outcomes.

1. First round offer rejection (c.f. Ochs and Roth (1989)). In LFE, the first period proposal of 1_1 and 2_1 is $(700, 300)$. This proposal is rejected by 2_2 and 3_2 because they fail to take into account that player-1 has absolute bargaining power in the last (third) period and that the pie will shrink to 600 in the next (second) period when they will have to make a counterproposal.[37]

2. First offers with near equal split or an offer equal to the second round pie (c.f. Neelin et al (1988); Guth and Tietz (1988); Ochs and Roth (1989)): 3_1, 4_1, and 5_1 propose $(580, 420)$ in the first period. 3_1, and 4_1 choose $(580, 420)$ because they cannot foresee that they will have absolute bargaining power in the last (third) period. In particular, 3_1 and 4_1 think that their payoff in the last period will be 180, and not 360. Thus 3_1, and 4_1 think that they must offer $600 - 180 = 420$ to player-2 in the first period to obtain immediate acceptance and the highest possible payoff. 5_1 chooses $(580, 420)$ despite having no foresight limitation. This is because 5_1 gets immediate acceptance from all player-2 types with this offer. If 5_1 were to offer less than 420 to player-2, his offer would be rejected by 1_2 and 2_2. Note that the prior probability on 1_2 and 2_2 ($\frac{2}{5}$ in this uniform distribution case) is crucial here. For example, if 5_1 knew that player-2 is the no-foresight-limitation type with a "high enough" probability, then 5_1 would propose the SPNE split of $(760, 240)$ because he would face rejection from only the limited-foresight player-2 types, but the probability of player-2 being a limited-foresight type would be low enough.

3. *Disadvantageous counter proposals* (c.f. Ochs and Roth (1989)). The LFE model

[37] 0_1's first round offer is also rejected. 0_1 overestimates his bargaining position in the first period due to limited foresight. Thus, 0_1 demands the whole first period pie. This demand is rejected by all player-2 types. These outcomes do not seem realistic. In applications, one can choose ρ to fit the data. So, in this Bargaining application it might be better to put 0 prior probability on 0_1.

generates disadvantageous counteroffers when 1_1 is matched with 3_2 or 2_1 is matched with 3_2. Player-types 1_1 and 2_1 make a proposal of $(700, 300)$ in the first period. However, 2_2 and 3_2 reject anything that gives them less than 420 because, due to their limited foresight, they think that all player-1 types will accept a proposal of $(180, 420)$ in the second period. However, once 3_2 moves forward and reaches the second period, he observes the full three-period Bargaining game. So, in the second period, 3_2 observes the absolute bargaining power of his opponent in the last (third) period. Thus, after rejecting a proposal of $(700, 300)$ in the first period, it is sequentially rational for 3_2 to make a disadvantageous counterproposal of $(360, 240)$ in the second period. A theoretical prediction of the LFE model is that this feature should disappear if we change the extensive form and make player-2 think about the acceptance/rejection decision simultaneously with the counterproposal decision. Thus, one should take great care in matching the specification of moves in the game to the foresight of the players.

4. Subgame consistency violation (c.f. Binmore et al (2002)):[38] Consider the two-period Bargaining game with the starting pie of 600 being tested separately and its data being compared to the data generated from the subgame consisting of the last two periods of the three-period Bargaining game. LFE shows that the results of the former may not match the data generated from the latter because these seemingly perfect information games may, in fact, be Interaction Games. In the three-period Bargaining game we consider, the outcome of the last two periods depends on what happened in the first period. For example, if the first proposal were $(700, 300)$, then player-1's type can be 1_1 or 2_1 with equal probability. These different player-1 types, have different optimal acceptance thresholds for the second period offer that they receive. So player-2's types,

[38] Binmore et al (2002) studied subgame consistency violation using the one-period subgame of a two-period Bargaining game. The LFE model cannot explain their findings because in the last period, limited-foresight has no bite. So all types are supposed to make "rational" choices in the last period, regardless of prior moves. However, if the subgame has more than one period, then we will show that LFE can explain subgame consistency violation.

3_2 and 4_2, update about player-1's type based on the first period proposal and adjust their optimal proposals in the second period. However, if two players are beginning a two-period Bargaining game, then their optimal choices only depend on their prior belief about the opponent's foresight, which may well be different from their updated belief going into period-two after observing the opponent's choice in the first period of a three-period Bargaining game.

Thus the LFE concept provides us several channels to explain several qualitative features of the data on Sequential Bargaining experiments. Fitting experimental data using this model is left as future work.

5.2 The Centipede Game

The Centipede game (Rosenthal (1981)) describes a situation in which two players alternately decide whether to *take* or *pass* a pile of money which increases whenever a player passes it to the opponent. Consider an S-staged Centipede game. First, player-1 decides whether to *take* or *pass* a pile of money; if the player moving at stage i decides to *take* at stage i then he gets a_i, the larger share of the existing pile of money, $a_i + b_i$, while his opponent gets b_i. If the player *passes*, the pile of money grows to $a_{i+1} + b_{i+1}$, that is, $a_i + b_i < a_{i+1} + b_{i+1}$ holds. But if the player *passes*, and his opponent *takes* in the next stage, he gets a payoff of b_{i+1}, and $b_{i+1} < a_i$ holds. However, if his opponent *passes* in stage-$(i+1)$, then the pile grows again and the player has a chance to *take* in stage-$(i+2)$ and achieve a payoff of a_{i+2}, which is strictly greater than a_i. In the last stage, if the player moving there *takes* then he gets a payoff of a_S, while his opponent gets b_S, where $a_{S-1} < b_S < a_S$ holds. If the player moving at the last stage chooses *pass* then his payoff is b_{S+1}, which is strictly less than a_S. The unique SPNE prediction is that the first player should *take* in the very first stage, regardless of the number of stages that the pile can be passed and grown. The logic is that in the last stage, as $a_S > b_{S+1}$ holds, so the player moving there should *take* ; but given this, one should *take* in the second-last stage, and this optimality of *taking* given one's opponent is going to *take*

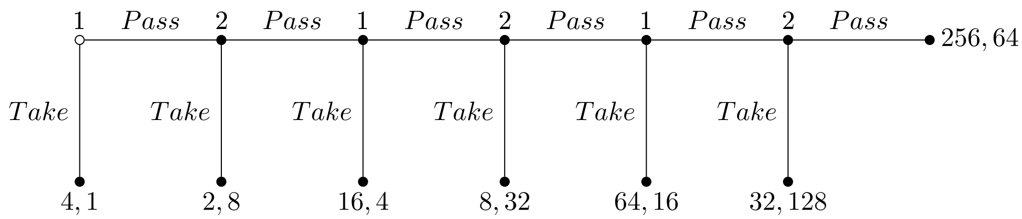

Figure 7. The Six-Staged Centipede Game

in the next stage continues inexorably backwards, and leads to the SPNE prediction: *take* in the first stage. This is unintuitive and various experiments, eg. McKelvey and Palfrey (1992, 1998) reject the SPNE prediction. Consider an S-staged Centipede game as Γ_0. We restrict our analysis to the Centipede games with the following, commonly used, payoff structure.

Definition 7: *An S-staged Centipede game is said to have the payoff structure P if for all $i \in \{1, .., S+1\}$:* (i) $b_i < b_{i+1} < a_i < b_{i+3} < a_{i+2}$ (ii) $a_i < \frac{b_{i+1}+a_{i+2}}{2}$ (iii) $\frac{a_i - b_{i+1}}{a_{i+2}-b_{i+1}} = \eta_i < \frac{1}{3}$.

The six staged Centipede game used by McKelvey and Palfrey (1992), depicted in Figure 7, also has the payoff structure P with $\eta_i = \frac{1}{7}$ for all i.[39] The condition $b_{i+1} < a_i < a_{i+2}$ follows from Γ_0 being a Centipede game. Conditions (ii) and (iii) of Definition 7 will be used in proving Proposition 6 below. An S-staged perfect information Centipede game Γ_0 generates an Interaction Game Γ with the player set $N = \{0_1,\ 1_1, ...,\ (S-1)_1, 0_2,\ 1_2, ...,\ (S-2)_2\}$.

The following proposition says that given a certain form of prior distribution over limited-foresight types, even with arbitrary positive total prior probability on limited-foresight types, all LFE outcomes entail *pass* being played with strictly positive probability by all foresight-types (including the no-foresight-limitation type) of both players, until the end stages of a Centipede game.

Proposition 6: *Consider an S-staged Centipede game Γ_0 with payoff structure P. Γ_0 generates an S-staged Interaction Game Γ. Let ρ, the prior distribution in Γ be such that*

[39]If a term, for example b_{i+3}, does not exist then any condition on that term is satisfied vacuously.

$$Pr(j_1) = Pr(k_2) = q \in (0,1] \; \forall \; j = 0,1,..,S-2 \text{ and } \forall \; k = 0,1,..,S-3,$$

$$\text{and} \sum_{j=0}^{S-1} Pr(j_1) = \sum_{k=0}^{S-2} Pr(k_2) = 1 \text{ hold.}$$

$$\text{So } Pr((S-1)_1) = 1-(S-1)q \text{ and } Pr((S-2)_2) = 1-(S-2)q, \text{ hold.}$$

Further, suppose the distribution on player-1's types is independent of the distribution on player-2's types. Then in any LFE of Γ*, all types of both players "pass" with strictly positive probability from stage-1 through stage-(S-3).*

The proof is given in the Appendix. This result depends on the following argument. First, player-types with total foresight strictly less than S always *pass* because according to the Curtailed Game they observe, the highest payoff for all types of either player occurs when both players *pass* until the end of that Curtailed Game. Next, in LFE, the player-type with no foresight limitation ((loosely) referred to as the "rational" player-type) pretends to be a low-foresight type and attains a higher payoff by *passing* until the end stages. The rational player-type can *pass* because, in LFE, his rational opponent reciprocates by *passing* with a high enough probability until the end stages of the game. The rational opponent reciprocates by *passing* because when the rational opponent observes *pass*, he believes with a high enough probability that he is playing against someone who will have strictly limited-foresight in the next stage, i.e., someone who will *pass* for sure. This analysis is almost parallel to the McKelvey and Palfrey (1992) model without the errors in actions, heterogeneous beliefs and learning components. In particular, if we have a ρ such that $Pr(0_1) = Pr(0_2) = 1-q$ and $Pr((S-1)_1) = Pr((S-2)_2) = q$ hold, then we can use McKelvey and Palfrey (1992) to characterize the unique LFE.[40] Both these analyses are in the same vein as the reputation

[40] The only difference would be that we would have to replace S by $S-1$ in their analysis as their altruist type (corresponding to 0_1, 0_2), who occurs with probability $(1-q)$ chooses *pass* in all stages. However, in

literature starting from Kreps, Milgrom, Roberts and Wilson (1982). The key takeaway from the Centipede game result is the applicability of the LFE in explaining multiple existing puzzles regarding perfect information games without using game specific "crazy" types.

6 Modeling Subjective Beliefs

The assumption of a common-knowledge prior distribution over foresight-levels across players simplifies the LFE model. This assumption implies that while player-types foresee different numbers of stages of the Interaction Game being played, each player-type has the same prior belief about the distribution over foresight-levels; and this prior belief is identical to the "true" distribution. The common prior assumption implies that we can use the same Curtailed Game to solve for the LFE actions of all player-types who have the same total foresight (recall that total foresight is defined as foresight-level plus stage-number of the move), even though their foresight-levels may be different. However, if we allow for subjective prior beliefs, that is, prior beliefs that may be different from the "true" prior distribution and vary based on the type/foresight-level of the player-type, then we introduce another dimension along which player-types have a subjective view of the Interaction Game (the first one being different foresight-levels).

In the subjective beliefs framework, at each move of each player-type, we have to construct a different Curtailed Game that models both the player-type's total foresight and his subjective belief about the prior distribution. In particular, we may not be able use the same Curtailed Game for player-types with the same total foresight if they have different prior beliefs. For example, consider again the Centipede game example in Figures 1-5. In Figure 4, we use $MCG(2)$ to solve for the LFE actions of both 1_1 at stage-one and 0_2 at stage-two, because their total foresights at their respective moves are the same, 2, and they both have the same prior beliefs: The prior distribution specified in the Interaction Game (Figure 2's notes). Without the common prior assumption, we will need different Curtailed Games for

the LFE model, even the lowest foresight types in our analysis, 0_1 and 0_2, *take* in the S^{th} stage.

1_1 at stage-one and 0_2 at stage-two. For 0_2 (respectively 1_1) we need a two-staged Curtailed Game with a prior distribution given by his subjective belief: ρ_{0_2} (respectively ρ_{1_1}). In this section, we will first describe the LFE for the case where the prior beliefs vary based on the foresight-level of the player: For example, ρ_{0_2} is different from ρ_{1_1}. It is straightforward to extend this model to the case where a player may have different types with the same foresight-level but different prior beliefs. That is, the case where the description of a type includes both foresight-level and prior beliefs. For example, the latter case would allow for a specification such that there can be two different types of player-2: Both have a foresight-level equal to 0, but the first has the prior beliefs given by $\rho_{0_2}(1)$, and the second has the prior beliefs given by $\rho_{0_2}(2)$, and $\rho_{0_2}(1) \neq \rho_{0_2}(2)$. We discuss such extensions at the end of this section.

We will first model the case of subjective prior beliefs where a player-type's foresight-level determines his prior beliefs. Consider an underlying sequential game with perfect information game, Γ_0. In this model, we will consider a player-type-specific corresponding Interaction Game, $\Gamma_{\rho_{t_i}} = \{N,\ H,\ \{\mathcal{I}(t_i)\}_{t_i \in N},\ P,\ A, \{u_{t_i}\}_{t_i \in N}, \rho_{t_i}\}$, for each player-type $t_i \in N$ and solve for his actions and beliefs using the Curtailed Games constructed from $\Gamma_{\rho_{t_i}}$. The Interaction Games $\Gamma_{\rho_{t_i}}$ differs across player-types as ρ_{t_i} is possibly different across $t_i \in N$.

We will work with the following two assumptions in what follows. Assumption 2 helps restrict the belief hierarchy of the player-types to their second-order beliefs. Assumption 3 is the subjective belief analogue of Assumption 1.

Assumption 2: Each player-type t_i believes that other player-types also have the same prior belief about the distribution over foresight-levels across players as her own subjective prior belief, ρ_{t_i}.

Assumption 3: For every player i and for every player-type t_i, her subjective prior belief about the distribution over foresight-levels across players, ρ_{t_i}, is such that conditional on t_i, the probability of the profile $(S - s_j)_{j \neq i}$ is strictly greater than 0. That is,

$$\text{Prob}((t_j \text{ has no foresight limitation})_{j \neq i} | \rho_{t_i}, t_i) = \text{Prob}((t_j = (S - s_j)_j)_{j \neq i} | \rho_{t_i}, t_i) > 0,$$

$$\forall i \in N_0, \ \forall t_i \in T_i.$$

Now we define an LFE with Subjective Beliefs (henceforth LFESB). Consider the "true" Interaction Game $\Gamma = \{N, \ H, \ \{\mathcal{I}(t_i)\}_{t_i \in N}, \ P, \ A, \{u_{t_i}\}_{t_i \in N}, \rho\}$ which has the "true" prior distribution denoted by ρ.[41] Let $N(\hat{\rho})$ be the subset of player-types who have the same subjective prior beliefs: $\hat{\rho}$. That is, $N(\hat{\rho}) \equiv \{t_i \in N\colon \rho_{t_i} = \hat{\rho}\}$. As the set of player-types, N, is finite, there is a finite set of possible subjective prior beliefs, $\hat{\rho}$, such that $N(\hat{\rho})$ is non-empty. The definition and construction of an LFESB for Γ works as follows. Consider all the different subjective prior beliefs that different player-types have. For each such subjective prior beliefs, say $\hat{\rho}$, construct $\Gamma_{\hat{\rho}}$ and solve for its LFE. If there are more than one LFE of $\Gamma_{\hat{\rho}}$, then arbitrarily select one LFE. Call it $(\pi^{\hat{\rho}}, \mu^{\hat{\rho}})$. To construct the LFESB strategy and belief for the player-types in $N(\hat{\rho})$, set $(\pi_{t_i}, \mu_{t_i}) = (\pi^{\hat{\rho}}_{t_i}, \mu^{\hat{\rho}}_{t_i})$ for all t_i in $N(\hat{\rho})$. Repeat these steps for all $\hat{\rho}$ such that $N(\hat{\rho})$ is non-empty to complete the construction of an LFESB. So the definition of LFESB is as follows.

Definition 8: (π, μ) *is a* ***Limited Foresight Equilibrium with Subjective Beliefs*** *of* Γ *if it is constructed as follows. For each subjective belief about the prior distribution* $\hat{\rho}$ *such that* $N(\hat{\rho}) \subset N$ *is non-empty, fix an LFE of* $\Gamma_{\hat{\rho}}$ *denoted by* $(\pi^{\hat{\rho}}, \mu^{\hat{\rho}})$. *For all* $t_i \in N(\hat{\rho})$, *set* $(\pi_{t_i}, \mu_{t_i}) = (\pi^{\hat{\rho}}_{t_i}, \mu^{\hat{\rho}}_{t_i})$.

Due to Proposition 1, it follows that an LFESB always exists for any finite Γ. Note that using $(\pi^{\hat{\rho}}, \mu^{\hat{\rho}})$ to construct the LFESB strategy and belief for all $t_i \in N(\hat{\rho})$ is a consistency

[41] By "true" we mean that ρ is the distribution over foresight-levels across players in the population (for a given underlying game). In a statistical exercise, ρ would be used by the researcher to fit observed data on the underlying sequential game with perfect information.

condition. In particular, this definition nests the common prior case when $\rho_{t_i} = \rho$ for all $t_i \in N$. The construction of an LFESB in Definition 8 implies that every player-type is optimizing using a move specific MCG where this MCG is constructed using ρ_{t_i} as the prior distribution over foresight-levels across players. The researcher can fit observed data using this model. The "free parameters" will be the "true" prior distribution ρ, and the set of subjective prior beliefs $(\rho_{t_i})_{t_i \in N}$. The latter will generate a LFESB strategy profile given by π. Together, ρ and π will imply a probability distribution over outcomes of the underlying sequential game with perfect information, which can be compared with data.

It is worth noting that the LFESB model can be extended to the case where a player may have different types with the same foresight-level but different prior beliefs. That is, the case where the description of a type includes both foresight-level and prior beliefs. This case is modeled as follows. We allow for $R_{t_i} < \infty$ different player-types corresponding to t_i. In the Interaction Game, the set of types corresponding to i, T_i, would now include $t_i(1)$, $t_i(2)$,..., $t_i(R_{t_i})$, instead of just $t_i \in T_i$. For each $r \in \{1, .., R_{t_i}\}$, $t_i(r)$ has a different subjective prior given by $\rho_{t_i(r)}$. The "true" distribution ρ would include these additional player-types $\{t_i(1)$, $t_i(2)$,..., $t_i(R_{t_i})\}$ in its support. Given Assumption 2, each player-type assumes that all other player-types have the same prior belief as his own. Thus, each of the other players'-types will not be able to differentiate among $\{t_i(1), t_i(2), ..., t_i(R_{t_i})\}$. Therefore, this model will be exactly like the LFESB setup where the LFE calculation needs to done for additional player-types with their own subjective prior beliefs.

Conclusion

This paper defines the Limited Foresight Equilibrium (LFE). The LFE is defined for general applicability in the class of finite sequential move games with perfect information. We model the case where players are interested in maximizing own payoff, but each player possesses one of different levels of foresight. Further, players are uncertain about their opponents' foresight. The LFE model nests the perfect information case. We prove the existence, upper hemi-continuity and other properties of LFE. These properties are: (a) The higher the foresight-level of a player, the better he can estimate his opponents' foresight-level. (b) The more moves any player-type observes, the better he becomes at guessing the opponent's foresight-level. (c) If a particular foresight-type is surprised by a sequence of moves impossible against lower types, he discovers that he is playing against some higher type, and must use his total foresight at that move to optimize. (d) The high foresight type must choose between revealing his type or pretending to be a low type. We show the applicability of LFE in two existing puzzles in the class of finite, two player alternate move games, namely, the Centipede game and the Sequential Bargaining game. In the Centipede game, LFE unleashes reputation effects, as in Kreps, Milgrom, Roberts and Wilson (1982), and McKelvey and Palfrey (1992), which lead to cooperative behavior even among players with no foresight limitation. In the Sequential Bargaining application, the properties of LFE help rationalize the disparate findings from the study of bargaining: That is, LFE produces outcomes that show (i) first round offer rejection (ii) first round offer of near equal split (iii) disadvantageous counterproposals (iv) subgame inconsistency. These LFE results for Sequential Bargaining are parallel to several qualitative results in different experimental studies on bargaining. Last, the LFE is extended to the case where limited-foresight types have subjective prior beliefs about the distribution over foresight-levels across players.

References

Agranov, M., E. Potamites, A. Schotter, and C. Tergiman (2012). Beliefs and endogenous cognitive levels: An experimental study. *Games and Economic Behavior 75*(2), 449–463.

Alaoui, L. and A. Penta (2016). Endogenous depth of reasoning. *The Review of Economic Studies 83*(4), 1297–1333.

Aumann, R. J. (1995). Backward induction and common knowledge of rationality. *Games and Economic Behavior 8*(1), 6–19.

Battigalli, P. (1997). On rationalizability in extensive games. *Journal of Economic Theory 74*(1), 40–61.

Ben-Porath, E. (1997). Rationality, nash equilibrium and backwards induction in perfect-information games. *The Review of Economic Studies 64*(1), 23–46.

Binmore, K. (1996). A note on backward induction. *Games and Economic Behavior 17*(1), 135–137.

Binmore, K., J. McCarthy, G. Ponti, L. Samuelson, and A. Shaked (2002). A backward induction experiment. *Journal of Economic theory 104*(1), 48–88.

Binmore, K., A. Shaked, and J. Sutton (1985). Testing noncooperative bargaining theory: A preliminary study. *The American Economic Review 75*(5), 1178–1180.

Bolton, G. E. (1991). A comparative model of bargaining: Theory and evidence. *The American Economic Review*, 1096–1136.

Bonanno, G. (2001). Branching time, perfect information games, and backward induction. *Games and Economic Behavior 36*(1), 57–73.

Brandenburger, A. and A. Friedenberg (2014). Are admissibility and backward induction consistent? *Working Paper*.

Camerer, C. F., T.-H. Ho, and J.-K. Chong (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, 861–898.

Costa-Gomes, M., V. P. Crawford, and B. Broseta (2001). Cognition and behavior in normal-form games: An experimental study. *Econometrica 69*(5), 1193–1235.

Costa-Gomes, M. A. and V. P. Crawford (2006). Cognition and behavior in two-person guessing games: An experimental study. *The American economic review 96*(5), 1737–1768.

Crawford, V. P., M. A. Costa-Gomes, and N. Iriberri (2013). Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. *Journal of Economic Literature 51*(1), 5–62.

Crawford, V. P. and N. Iriberri (2007a). Fatal attraction: Salience, naivete, and sophistication in experimental hide-and-seek games. *The American Economic Review 97*(5), 1731–1750.

Crawford, V. P. and N. Iriberri (2007b). Level-k auctions: Can a nonequilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions? *Econometrica 75*(6), 1721–1770.

Gill, D. and V. Prowse (2016). Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis. *Journal of Political Economy 124*(6), 1619–1676.

Guth, W. and R. Tietz (1988). Ultimatum bargaining for a shrinking cake-an experimental analysis. In *Bounded Rational Behavior in Experimental Games and Markets*, pp. 111–128. Springer.

Guth, W. and R. Tietz (1990). Ultimatum bargaining behavior: A survey and comparison of experimental results. *Journal of Economic Psychology 11*(3), 417–449.

Ho, T.-H., C. Camerer, and K. Weigelt (1998). Iterated dominance and iterated best response in experimental p-beauty contests. *The American Economic Review 88*(4), 947–969.

Ho, T.-H. and X. Su (2013). A dynamic level-k model in sequential games. *Management Science 59*(2), 452–469.

Jeheil, P. (1995). Limited horizon forecast in repeated alternate games. *Journal of Economic Theory 67*(2), 497–519.

Jehiel, P. (1998a). Learning to play limited forecast equilibria. *Games and Economic Behavior 22*(2), 274–298.

Jehiel, P. (1998b). Repeated games and limited forecasting. *European Economic Review 42*(3), 543–551.

Jehiel, P. (2001). Limited foresight may force cooperation. *The Review of Economic Studies 68*(2), 369–391.

Jehiel, P. (2005). Analogy-based expectation equilibrium. *Journal of Economic theory 123*(2), 81–104.

Jehiel, P. and D. Samet (2007). Valuation equilibrium. *Theoretical Economics 2*(2), 163–185.

Johnson, E. J., C. Camerer, S. Sen, and T. Rymon (2002). Detecting failures of backward induction: Monitoring information search in sequential bargaining. *Journal of Economic Theory 104*(1), 16–47.

Kawagoe, T. and H. Takizawa (2012). Level-k analysis of experimental centipede games. *Journal Of Economic Behavior & Organization 82*(2), 548–566.

Ke, S. (2017). Boundedly rational backward induction. *Working Paper*.

Kreps, D. M., P. Milgrom, J. Roberts, and R. Wilson (1982). Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic theory 27*(2), 245–252.

Kreps, D. M. and R. Wilson (1982a). Reputation and imperfect information. *Journal of economic theory 27*(2), 253–279.

Kreps, D. M. and R. Wilson (1982b). Sequential equilibria. *Econometrica: Journal of the Econometric Society*, 863–894.

Levin, D. and L. Zhang (2016). Bridging level-k to nash equilibrium. *Working Paper*.

Levitt, S. D., J. A. List, and E. Sally (2011). Checkmate: Exploring backward induction among chess players. *The American Economic Review 101*(2), 975–990.

Mantovani, M. (2014). Limited backward induction: Foresight and behavior in sequential games. *Working Paper*.

McKelvey, R. D. and T. R. Palfrey (1992). An experimental study of the centipede game. *Econometrica: Journal of the Econometric Society*, 803–836.

McKelvey, R. D. and T. R. Palfrey (1998). Quantal response equilibria for extensive form games. *Experimental economics 1*(1), 9–41.

Milgrom, P. and J. Roberts (1982). Predation, reputation, and entry deterrence. *Journal of economic theory 27*(2), 280–312.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *The American Economic Review 85*(5), 1313–1326.

Neelin, J., H. Sonnenschein, and M. Spiegel (1988). A further test of noncooperative bargaining theory: Comment. *The American Economic Review 78*(4), 824–836.

Ochs, J. and A. E. Roth (1989). An experimental study of sequential bargaining. *The American Economic Review*, 355–384.

Osborne, M. J. and A. Rubinstein (1994). *A course in game theory*. MIT press.

Palacios-Huerta, I. and O. Volij (2009). Field centipedes. *The American Economic Review 99*(4), 1619–1635.

Rampal, J. (2017). Opponent's foresight and optimal choices. *Working Paper*. Available at: https://sites.google.com/site/jeevantrampalecon/Research.

Reynolds, R. I. (1992). Recognition of expertise in chess players. *The American journal of psychology*, 409–415.

Roomets, A. (2010). On limited foresight in games. *Working Paper*.

Rosenthal, R. W. (1981). Games of perfect information, predatory pricing and the chain-store paradox. *Journal of Economic theory 25*(1), 92–100.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica: Journal of the Econometric Society*, 97–109.

Slonim, R. L. (2005). Competing against experienced and inexperienced players. *Experimental Economics 8*(1), 55–75.

Stahl, D. O. (1996). Boundedly rational rule learning in a guessing game. *Games and Economic Behavior 16*(2), 303–330.

Stahl, D. O. and P. W. Wilson (1995). On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior 10*(1), 218–254.

Stahl, I. (1972). Bargaining theory. Economic Research Institute, Stockholm School of Economics, Stockholm.

Appendix

Proof of Proposition 1(a). *(Existence of LFE)*: Consider an arbitrary finite Interaction Game Γ. The $CG(1)$ derived from Γ is also finite. Due to Proposition 1 of Kreps and Wilson (1982b), there exists a SE of $CG(1)$. We can select an arbitrary $SE(1)$ of $CG(1)$ to construct $MCG(2)$. $MCG(2)$ is also finite. Thus the $SE(2)$ of $MCG(2)$ also exists. Proceeding thus, given the existence of SE for each of $CG(1)$, $MCG(2)$,...,$MCG(n-1)$, we can construct $MCG(n)$ in step n of Definition 5. As $MCG(n)$ is finite, there exists a SE of $MCG(n)$. As this holds for $n = 2, ..., S$, each of the steps in Definition 5 is well defined. Thus, there exists at least one LFE of Γ. *Q.E.D.*

Proof of Proposition 1(b). *(Upper hemi-continuity of LFE)*: Consider an arbitrary S-staged extensive form, $\{N,\ H,\ \{\mathcal{I}(t_i)\}_{t_i \in N},\ P,\ A\}$, generated from Γ_0 using the $Seq(.)$ correspondence. Let the correspondence $f\ :\ \Delta T \times R^N \implies \Pi \times \mathcal{M}$ be the set valued function, mapping a tuple of initial conditions, i.e. a tuple comprising of the prior distribution and payoffs,(ρ, u), to the set containing all the LFE assessments of the Interaction Game, Γ, so defined. An element of the set $f(\rho, u)$ is an LFE of Γ, denoted as (π, μ). Fix a sequence (ρ_k, u_k) such that $(\rho_k, u_k) \to (\rho, u)$ and an associated sequence $(\pi_k, \mu_k) \in f(\rho_k, u_k)$, such that $(\pi_k, \mu_k) \to (\pi, \mu)$. To show upper hemi-continuity, we need to show that $(\pi, \mu) \in f(\rho, u)$.

Given an arbitrary extensive form $\{N',\ H',\ \{\mathcal{I}'(t_i)\}_{t_i \in N'}\ P',\ A'\}$, let $\Psi :\ \Delta T \times R^{N'} \implies \Sigma \times B$ be the upper hemi-continuous (by Proposition 2 of Kreps and Wilson (1982b)) correspondence mapping the common prior and payoffs, (ρ', u'), to the set $\Psi(\rho', u')$, which contains all the Sequential Equilibria assessments, (σ, b), of the game so defined. Let $\pi(\mathcal{I})$, $\mu(\mathcal{I})$ denote the vectors π and μ restricted to the coordinates corresponding to the information sets contained in $\mathcal{I}$. Let $f(\rho, u)(\mathcal{I})$ also represent each LFE assessment in $f(\rho, u)$ restricted to $\mathcal{I}$. We will now prove upper hemi-continuity by induction.

Step 1. We show that $(\pi(D^1), \mu(D^1)) \in f(\rho, u)(D^1)$. Consider $CG(1) \equiv MCG(1)$. Corresponding to (ρ_k, u_k) we have (ρ_k^1, u_k^1) for each element of the sequence $k = 1, 2, ...$ The superscript denotes the length of the CG. The construction of u_k^1 using the $\frac{min+max}{2}$ curtailment method was described earlier. As the function which maps a finite set of real numbers to their $\frac{min+max}{2}$ is a continuous function, $u_k \to u$ implies $u_k^1 \to u^1$. Also, $\rho_k^1 = \rho_k$ holds for all k, so $(\rho_k, u_k) \to (\rho, u)$ implies $(\rho_k^1, u_k^1) \to (\rho^1, u^1)$. Note that for each k in the sequence, $(\pi_k(D^1),\ \mu_k(D^1)) = (\sigma_k^1(D^1),\ b_k^1(D^1))$ holds for some $(\sigma_k^1(D^1),\ b_k^1(D^1)) \in \Psi(\rho_k^1, u_k^1)(D^1)$. We are given that $(\pi_k(D^1), \mu_k(D^1))$ converges to $(\pi(D^1), \mu(D^1))$. Thus, $(\sigma_k^1(D^1)$,$b_k^1(D^1)) = (\pi_k(D^1), \mu_k(D^1))$ is a convergent sequence. We know that $\Psi(.)$ is upper hemi-continuous. Thus if $(\sigma_k^1(D^1)$,$b_k^1(D^1)) \to (\sigma^1(D^1),\ b^1(D^1))$ then $(\sigma^1(D^1),\ b^1(D^1)) \in \Psi(\rho^1, u^1)(D^1)$. Given that $(\pi_k(D^1)$,$\mu_k(D^1)) \to (\pi(D^1),\ \mu(D^1))$, and given that $(\pi_k(D^1),$

$\mu_k(D^1)) = (\sigma_k^1(D^1),\ b_k^1(D^1)) \to (\sigma^1(D^1), b^1(D^1))$ it follows from the uniqueness of a limit that $(\sigma^1(D^1),\ b^1(D^1)) = (\pi(D^1),\ \mu(D^1)) \in \Psi(\rho^1, u^1)(D^1) \subset f(\rho,\mu)(D^1)$. Therefore $(\pi(D^1)$, $\mu(D^1)) \in f(\rho,\mu)(D^1)$.

Step n**.** Consider $MCG(n)$, where $n \in \{2,..,S\}$. Let $(\pi(\bigcup_{i=1}^{i=n-1} D^i),\ \mu(\bigcup_{i=1}^{i=n-1} D^i))$ $\in f(\rho,u)(\bigcup_{i=1}^{i=n-1} D^i)$. We will show that $(\pi(\bigcup_{i=1}^{i=n} D^i),\ \mu(\bigcup_{i=1}^{i=n} D^i)) \in f(\rho,u)(\bigcup_{i=1}^{i=n} D^i)$. Given step 1, this will complete the proof. Corresponding to (ρ_k, u_k) we have (ρ_k^n, u_k^n) for each $k = 1, 2, \ldots$. Using $\pi_k(\bigcup_{i=1}^{i=n-1} D^i)$ and ρ_k we generate ρ_k^n as detailed earlier in the construction of $MCG(n)$. By continuity, $u_k^n \to u^n$ holds. As $\pi_k(\bigcup_{i=1}^{i=n-1} D^i) \to \pi(\bigcup_{i=1}^{i=n-1} D^i)$ by assumption, thus: (i) $(\rho_k^n, u_k^n) \to (\rho^n, u^n)$ and (ii) it will suffice to show $(\pi(D^n),\ \mu(D^n)) \in f(\rho,u)(D^n)$. Now note that for each k, $(\pi_k(D^n),\ \mu_k(D^n)) = (\sigma_k^n(D^n),\ b_k^n(D^n))$ for some $(\sigma_k^n(D^n),\ b_k^n(D^n))$ $\in \Psi(\rho_k^n, u_k^n)(D^n)$. We are given that $(\pi_k(D^n), \mu_k(D^n))$ converges to $(\pi(D^n), \mu(D^n))$. Thus, $(\sigma_k^n(D^n), b_k^n(D^n)) = (\pi_k(D^n), \mu_k(D^n))$ is a convergent sequence. We know that $\Psi(.)$ is upper hemi-continuous. Thus, if $(\sigma_k^n(D^n), b_k^n(D^n)) \to (\sigma^n(D^n),\ b^n(D^n))$, then $(\sigma^n(D^n),\ b^n(D^n))$ $\in \Psi(\rho^n, u^n)(D^n)$. Given that $(\pi_k(D^n), \mu_k(D^n)) \to (\pi(D^n), \mu(D^n))$, and given that

$$(\pi_k(D^n), \mu_k(D^n)) = (\sigma_k^n(D^n),\ b_k^n(D^n)) \to (\sigma^n(D^n),\ b^n(D^n)),\ \text{holds},$$

by the uniqueness of a limit, it follows that

$$(\sigma^n(D^n), b^n(D^n)) = (\pi(D^n), \mu(D^n)) \in \Psi(\rho^n, u^n)(D^n) \subset f(\rho,\mu)(D^n).$$

Therefore, $(\pi(D^n),\ \mu(D^n)) \in f(\rho,u)(D^n)$. *Q.E.D.*

Proof of Proposition 2. Suppose the precedent of Proposition 2 holds and h_0 is the sequence of actions of the underlying sequential game with perfect information that generates the corresponding information sets $I(t_i)$ and $I(t_i')$ of $MCG(n)$. Then

$$b^n((t_{-i}, h_0)|\ t_i,\ I(t_i),\ \sigma^n, \rho^n) = \frac{Pr((t_i, t_{-i}, h_0)|\ \sigma^n, \rho^n)}{\sum_{t_{-i} \in T_{-i}} [Pr((t_i, t_{-i}, h_0)|\ \sigma^n, \rho^n)]} \tag{5}$$

We have to show that

$$\frac{Pr((t_i', t_{-i}, h_0)|\ \sigma^n, \rho^n)}{\sum_{t_{-i}\in T_{-i}}[Pr((t_i', t_{-i}, h_0)|\ \sigma^n, \rho^n)]} = \frac{Pr((t_i, t_{-i}, h_0)|\ \sigma^n, \rho^n)}{\sum_{t_{-i}\in T_{-i}}[Pr((t_i, t_{-i}, h_0)|\ \sigma^n, \rho^n)]} \text{ holds.} \tag{6}$$

The aim of the proof is to show that the right hand side (RHS) of (5) does not depend on t_i (the player-type), it only depends on i (the player). Without loss of generality, consider an action sequence $h_0 = (a_0^k)_{k=1,...,K}$. Let the information set containing a history (t, h_0) be denoted as $I(t, h_0)$. Using the independence of ρ we get the following.

$$\begin{aligned} Pr((t_i, t_{-i}, h_0)|\ \rho^n, \sigma^n) &= Pr(t_i|\rho).Pr(t_{-i}|\rho).Pr(a_0^1|I(t), \rho^n, \sigma^n) \\ .Pr(a_0^2|I(t, a_0^1), \rho^n, \sigma^n) &\ \ ...\ \ Pr(a_0^K|I(t, (a_0^k)_{k=1,...,K-1}), \rho^n, \sigma^n). \end{aligned} \tag{7}$$

Let (t, a_0^0) denote the stage-1 move immediately following Nature's selection of the player-type combination: t. Define a subsequence/subhistory of (t, h_0) as $(t, (a_0^k)_{k=0,...,r-1})$, such that $r \in \{1, .., K\}$ and there exists a unique sequence of actions $(a_0^k)_{k=r,...,K}$ such that

$$(t, (a_0^k)_{k=0,...,r-1}, (a_0^k)_{k=r,...,K}) = (t, h_0) \text{ holds.}$$

For any (t_i, t_{-i}, h_0) such that $t_{-i} \in T_{-i}$, let $R(i)$ be the collection of stage numbers of $MCG(n)$, $r(i) \in \{1, .., K\}$, such that t_i moves at the subhistory $((t_i, t_{-i}),(a_0^k)_{k=0,...,r(i)-1})$ of (t_i, t_{-i}, h_0). Further, let $R^c(i)$ be the stage numbers where some t_j moves, such that $j \neq i$ holds. That is, $R^c(i) = [\{1, .., K\} - R(i)]$. Note that, by the construction of Γ, the sets $R(i)$ and $R^c(i)$ do not depend on t_i or t_{-i}; they only depend on i and the h_0 component of (t, h_0). Recall that in any $MCG(n)$, the information sets only reflect the uncertainty about the opponents'-types, (t_{-i}). Therefore, given an action sequence h_0, different profiles of opponents'-types playing h_0 only implies different nodes in the same information set. So for any $t_{-i}, t_{-i}' \in T_{-i}$, $I((t_i, t_{-i}), (a_0^k)_{k=0,...,r(i)-1}) = I((t_i, t_{-i}'), (a_0^k)_{k=0,...,r(i)-1})$ holds by construction. Thus, we can rewrite (7) as:

$$\begin{aligned} Pr((t_i\ ,\ t_{-i},\ h_0)|\ \sigma^n,\rho^n) &= Pr(t_{-i}|\rho)\Pi_{s\in R^c(i)}[Pr(a_0^s|\ I(t,\ (a_0^k)_{k=0,\dots,s-1}),\rho^n,\sigma^n)] \\ &\times\ Pr(t_i|\rho)\Pi_{r\in R(i)}[Pr(a_0^r|\ I(t,\ (a_0^k)_{k=0,\dots,r-1}),\rho^n,\sigma^n)]. \end{aligned} \tag{8}$$

Equation (8) implies that we can cancel $Pr(t_i|\rho)\Pi_{r\in R(t_i)}[Pr(a_0^r|\ I(t,\ (a_0^k)_{k=0,\dots,r-1}),\rho^n,\sigma^n)]$ from both the numerator and denominator of both the RHS and LHS (left hand side) of (5). So the RHS of (5) can be written as

$$\frac{Pr((t_i,\ t_{-i},\ h_0)|\ \sigma^n,\rho^n)}{\sum_{t_{-i}\in T_{-i}}[Pr((t_i,\ t_{-i},\ h_0)|\ \sigma^n,\rho^n)]} = \frac{Pr(t_{-i}|\rho)\Pi_{s\in R^c(i)}[Pr(a_0^s|\ I(t,\ (a_0^k)_{k=0,\dots,s-1}),\rho^n,\sigma^n)]}{\sum_{t_{-i}\in T_{-i}} Pr(t_{-i}|\rho)\Pi_{s\in R^c(i)}[Pr(a_0^s|\ I(t,(a_0^k)_{k=0,\dots,s-1}),\rho^n,\sigma^n)]} \tag{9}$$

The proof will be complete if we show that the RHS of (9) does not depend on the type of player i: t_i. First note that the term $Pr(t_{-i}|\rho)$ does not depend on t_i. Next, as argued before, $R^c(i)$ also does not depend on t_i. Last, notice that for any $s\in R^c(i)$, the term $\sigma^n(a_0^s|\ I(t,\ (a_0^k)_{k=0,\dots,s-1}),\rho^n)$ does not depend on t_i. To see this, suppose the player-type moving at $(t,\ (a_0^k)_{k=0,\dots,s-1})$ is t_j. Then, for different t_i, given the preceding action sequence $(a_0^k)_{k=0,\dots,s-1}$, all the nodes $((t_i,\ t_j,\ t_{-(i,j)}),\ (a_0^k)_{k=0,\dots,s-1})$ where $t_i\in T_i$, are in the same information set for t_j: $I(t,\ (a_0^k)_{k=0,\dots,s-1}))$. Therefore, t_j must choose the same probability of the action a_0^s, regardless of i's type. Therefore, for the term on the RHS of (9), both the numerator and all the terms in the denominator are independent of $t_i\in T_i$. Thus, the RHS of (5) does not depend on t_i, and remains constant across $t_i, t_i'\in T_i$ for corresponding information sets, which implies that (6) holds. *Q.E.D.*

Proof of corollary to Proposition 2. Let $U_{t_i}^n(\sigma^n|\rho^n,b^n,I(t_i))$ denote the expected payoff of t_i from the probability distribution on the terminal histories of $MCG(n)$ implied by (σ^n,ρ^n,b^n), conditional on reaching the information set $I(t_i)$. By Proposition 2, if

$Seq^{-1}(I(t_i)) = Seq^{-1}(I(t_i'))$, then $b^n(I(t_i)) = b^n(I(t_i'))$. Further other players' ($j \neq i$) types cannot choose different actions for different types of player i. Thus, t_i and t_i' face the same strategy profile $((\sigma^n_{t_j})_{t_j \in T_j})_{j \neq i}$. In other words, $\sigma^n_{-t_i}$ is identical to $\sigma^n_{-t_i'}$ for the purpose of payoff calculation. Let $\mathcal{I}^{h_0}(t_i)$ denote all the information sets of t_i which follow the action sequence h_0. Let $\mathcal{I}^{h_0}(t_i')$ denote the corresponding information sets of t_i'. Let $I^{h_0}(t_i)$ and $I^{h_0}(t_i')$ be arbitrary corresponding information sets in $\mathcal{I}^{h_0}(t_i)$ and $\mathcal{I}^{h_0}(t_i')$, respectively. The arguments above imply that $U_{t_i}(s|\sigma^n_{-t_i}, b^n, I^{h_0}(t_i)) = U_{t_i'}(s|\sigma^n_{-t_i'}, b^n, I^{h_0}(t_i'))$ holds for all strategies $s(.)$ at corresponding information sets $I^{h_0}(t_i)$ and $I^{h_0}(t_i')$. Let $\sigma^n_{t_i'}$ denote the strategy identical to $\sigma^n_{t_i}$ (the only difference being that the former is played by t_i', the latter by t_i). Thus, the precedent of the corollary to Proposition 2 implies that

$$U_{t_i'}(\sigma^n_{t_i'}|\sigma^n_{-t_i'}, b^n, \rho^n, I^{h_0}(t_i')) > U_{t_i'}(s^n_{t_i'}|\sigma^n_{-t_i'}, b^n, \rho^n, I^{h_0}(t_i')) \text{ holds,}$$

for all $I^{h_0}(t_i') \in \mathcal{I}^{h_0}(t_i')$, and for all possible strategies $s^n_{t_i'}(\mathcal{I}^{h_0}(t_i'))$ over $\mathcal{I}^{h_0}(t_i')$. Given that $\sigma^n_{t_i}$ is t_i's strict best response at all $I^{h_0}(t_i) \in \mathcal{I}^{h_0}(t_i)$, and given that σ^n is a SE strategy profile, sequential rationality implies that $\sigma^n_{t_i'}(\mathcal{I}^{h_0}(t_i')) = \sigma^n_{t_i}(\mathcal{I}^{h_0}(t_i))$ must hold when $Seq^{-1}(\mathcal{I}^{h_0}(t_i)) = Seq^{-1}(\mathcal{I}^{h_0}(t_i'))$ holds. *Q.E.D.*

Proof of Proposition 3. Consider an arbitrary LFE, (π, μ), of Γ. Consider an arbitrary stage-K information set of Γ: I. Suppose the player-type moving at I is t_i, i.e., $P(I) = t_i$ holds. So t_i's total foresight at I is $(K + t_i)$. Let $(K + t_i) \leq S$.[42] Let h be an arbitrary node of I such that $h \in L(I)$. Let, without loss of generality, h be of the form $h = (t,\ (a_0^k)_{k=1,\ldots,s-1})$. The LFE belief $\mu_{t_i}(h|\ L(I))$ (corresponding to the LFE (π, μ)) is derived using $MCG(K + t_i)$ at step $(K + t_i)$ of the construction of an LFE specified in Definition 5. To construct $MCG(K + t_i)$, we need to complete steps 1 through $(K + t_i - 1)$ of Definition 5. In constructing $MCG(K + t_i)$, for any such $h \in L(I)$, using the definition of $L(I)$, it follows that for all subsequences of h of the form $\hat{h} = (t,\ (a_0^k)_{k=0,\ldots,r})$ such that $r \leq (K - 1)$ holds,

[42] The argument for $(K + t_i) > S$ is very similar.

$mP^{(K+t_i)}(\hat{h}) = \textit{Nature}$ holds because some MCG shorter than $MCG(K+t_i)$ is decisive for $\hat{h}$. That is, $\hat{h} \in \bigcup_{n=1}^{K+t_i-1} D^n$ holds. We know the LFE actions, $\pi(\bigcup_{n=1}^{K+t_i-1} D^n)$, by steps 1 though $(K+t_i-1)$ of the construction of an LFE. Further, these LFE actions are considered as Nature's moves in $MCG(K+t_i)$. Thus, in constructing $MCG(K+t_i)$, we set $\rho^{K+t_i}(\hat{h}) = \pi(\hat{h})$ for all subsequences $\hat{h}$ of each $h \in L(I)$. As $\mu_{t_i}(h|\ L(I))$ is calculated using the SE of $MCG(K+t_i)$, $\mu_{t_i}(h|\ L(I))$ is calculated using the Bayes' rule (wherever possible) and ρ^{K+t_i}, or equivalently, using Bayes' rule and (π, ρ). The equivalence follows because, by construction, ρ^{K+t_i} is identical to (π, ρ) wherever ρ^{K+t_i} is defined. Therefore, equation (3) holds wherever $Prob(L(I) \mid \rho^{K+t_i}) > 0$ holds. *Q.E.D.*

Proof of Remark 2. Suppose $t_i' > t_i$ holds. Let $h = ((t_i, t_{-i}),\ (a_0^k)_{k=1,\ldots,K-1}) \in L(I(t_i))$. We will show that $h' = ((t_i', t_{-i}),\ (a_0^k)_{k=1,\ldots,K-1}) \in L(I(t_i'))$ to complete the proof. Define $(t, a_0^0) \equiv t$. All subsequences of h can be written as $((t_i, t_{-i}),\ (a_0^k)_{k=0,\ldots,r-1})$ for some $r < K$. Fix one such arbitrary subsequence of h: $\hat{h} = ((t_i, t_{-i}),\ (a_0^k)_{k=0,\ldots,r-1})$. As $h \in L(I(t_i))$, it must be the case that if $P(\hat{h}) = t_j$, then $r + t_j < K + t_i$ holds. By the construction of Γ using the $Seq(.)$ function, we must have that for the same r, the subsequence of h' given by $\hat{h}' = ((t_i', t_{-i}),\ (a_0^k)_{k=0,\ldots,r-1})$ is such that $P(\hat{h}') = t_j$. Given that $t_i < t_i'$ holds, we must have that $r + t_j < K + t_i < K + t_i'$ holds. As the choice of $r < K$ was arbitrary, it follows that for all subsequences of h', the player-type moving at that subsequence has strictly lesser total foresight than t_i' at stage-K. Thus, if $h \in L(I(t_i))$ then $h' \in L(I(t_i'))$. *Q.E.D.*

Proof of Proposition 5. The LFE beliefs of t_i at $I(t_i)$ must be derived from some SE of $MCG(K+t_i)$. So fix an arbitrary SE of $MCG(K+t_i)$: $(\sigma^{K+t_i}, b^{K+t_i})$. By the LFE definition, all the subsequences preceding the nodes in $L(I(t_i))$ have Nature as the player moving there and the actions taken by Nature (given by ρ^{K+t_i}) are common knowledge in $MCG(K+t_i)$. Thus, the probability of reaching $I(t_i)$ via only Nature's moves, can be calculated using ρ^{K+t_i}

by t_i at $I(t_i)$. Thus, if $Pr(L(I(t_i))|\ \rho^{K+t_i}) = 0$ and $Pr(I(t_i))|\ \sigma^{K+t_i}, \rho^{s+t_i}) > 0$, then

$$b_{t_i}^{K+t_i}(L(I(t_i)) \mid I(t_i)) = \frac{Pr(L(I(t_i))|\ \rho^{K+t_i})}{Pr(I(t_i))|\ \sigma^{K+t_i}, \rho^{K+t_i})} = 0.$$

As $\mu_{t_i}(L(I(t_i)) \mid I(t_i)) = b_{t_i}^{K+t_i}(L(I(t_i)) \mid I(t_i))$, we have $\mu_{t_i}(L(I(t_i)) \mid I(t_i)) = 0$. Further, as

$$\mu_{t_i}([L(I(t_i))]^c \mid I(t_i)) + \mu_{t_i}(L(I(t_i)) \mid I(t_i)) = 1 \text{ holds,}$$

$$\mu_{t_i}([L(I(t_i))]^c \mid I(t_i)) = 1 \text{ holds.}$$

However, if $Pr(I(t_i))|\ \sigma^{K+t_i}, \rho^{K+t_i}) = 0$, then we cannot use the previous argument. In this case, by the definition of a SE, there must exist a consistent sequence $(\sigma^{K+t_i}, b^{K+t_i})_m$ such that $(\sigma^{K+t_i})_m$ is a totally mixed strategy profile in $MCG(K+t_i)$ and $(\sigma^{K+t_i}, b^{K+t_i})_m \to (\sigma^{K+t_i}, b^{K+t_i})$ as $m \to \infty$. For any such sequence, $Pr(I(t_i)|\ (\sigma^{K+t_i})_m, \rho^{K+t_i}) > 0$ holds for every m because of Assumption 1 and because $(\sigma^{K+t_i})_m$ is a totally mixed strategy profile.[43] Further, we are given that $Prob(L(I(t_i))|\ \rho^{K+t_i}) = 0$ holds. So it must hold that

$$(b_{t_i}^{K+t_i}(L(I(t_i)) \mid I(t_i)))_m = \frac{Prob(L(I(t_i))|\ \rho^{K+t_i})}{Prob(I(t_i)|\ (\sigma^{K+t_i})_m, \rho^{K+t_i})} = 0\ \forall m = 1, 2, ...$$

Thus $(b_{t_i}^{K+t_i}[L(I(t_i)]^c \mid I(t_i)))_m = 1$ for each m in the sequence. So for any SE $(\sigma^{K+t_i}, b^{K+t_i})$ of $MCG(K+t_i)$ we must have $b_{t_i}^{K+t_i}(L(I(t_i)) \mid I(t_i)) = 0$ and $b_{t_i}^{K+t_i}([L(I(t_i))]^c \mid I(t_i)) = 1$. Thus, by the definition of an LFE, we must have that $\mu_{t_i}(L(I(t_i)) \mid I(t_i)) = 0$ and $\mu_{t_i}([L(I(t_i))]^c \mid I(t_i)) = 1$. *Q.E.D.*

LFE calculation for the 3-period bargaining game. We assume the prior to be such that $Pr(t_1) = \frac{1}{6}$ for $t_1 \in \{0_1, 1_1, 2_1, 3_1, 4_1, 5_1\}$, and independently, $Pr(t_2) = \frac{1}{5}$ for $t_2 \in \{0_2, 1_2, 2_2, 3_2, 4_2\}$. Recall that every period has two stages: Proposal stage followed by the accept/reject decision stage. Let x_1 (respectively x_3) denote the demand of the first mover

[43] And because we are assuming that $I(t_i)$ is a non-trivial information set.

(player-1; P1 for short) for himself in the first stage (fifth stage), when the period number is one (three) and the size of pie is 1000 (respectively 360). Thus, $(1000 - x_1)$ (respectively $(360 - x_3)$) is the share of the first-stage (fifth-stage) pie offered to player 2 (P2 for short). The offer of P2 to P1 in the third-stage (period number two) is denoted as y_2. Thus $(600 - y_2)$ is the share of the third-stage pie demanded by P2 for himself. We summarize the LFE strategies in Table 1. $\overline{X_1}$ (respectively $\overline{X_3}$) denotes the maximum share of P1 out of the first (third) period pie, such that $(1000 - \overline{X_1})$ (respectively $(360 - \overline{X_3})$) is acceptable to P2 in period one (period three). The minimum share of the second-period pie offered by P2 to P1, such that it is acceptable to P1, is denoted as $\underline{Y_2}$. As per the definition of LFE, we construct the LFE starting with the SE of $CG(1)$. In what follows, we specify the SE and LFE beliefs only when needed to determine optimal actions.

Step 1: In $CG(1)$, as per the curtailment rule, the payoff of P1-types is increasing in x_1. Thus, in the unique $SE(1)$ of $CG(1)$, all P1-types choose $x_1 = 1000$ regardless of beliefs. We have solved for the LFE action at D^1, which only contains the information set of 0_1 at stage-I. Thus the LFE action of 0_1 at stage-I is $x_1 = 1000$.

Step 2: Fix 0_1's move at stage-I as Nature's move in $CG(2)$ to generate $MCG(2)$. The curtailment rule implies that the payoff of P2-types from rejecting P1's stage-I offer is 300. Thus, in the unique $SE(2)$[44] of $MCG(2)$, all P2 types at stage-II *accept* if $x_1 \leq 700$, regardless of belief on P1-types. Thus, in $SE(2)$, all P1-types other than 0_1 choose $x_1 = 700$, regardless of beliefs. We note the $SE(2)$ actions at D^2. This gives us the LFE actions at D^2, which contains the information sets of 1_1 and 0_2, at stage-I and stage-II, respectively.

Step 3: Fix the LFE actions at D^1 and D^2 as Nature's moves to convert $CG(2)$ to $MCG(3)$. The curtailment rule implies that P2-types' stage-III payoff is decreasing in y_2. Thus, in the unique $SE(3)$ of $MCG(3)$, all P2-types at stage-III choose $y_2 = 0$, irrespective of belief.

[44] This uniqueness is only of the SE strategy profile, not necessarily the belief profile. In what follows, we only consider the uniqueness of the SE strategy profile.

Thus, the stage-II $SE(3)$ action for all P2 types is to *accept* if $x_1 \leq 700$, regardless of belief on P1-types. Thus, in $SE(3)$, all P1-types propose $x_1 = 700$, regardless of beliefs. We note SE actions at D^3. This gives us the LFE actions at D^3, which contains the information sets of 2_1, 1_2, and 0_2, at stage-I, stage-II, and stage-III, respectively.

Step 4: Fix the LFE actions at $\bigcup_{n=1}^{3} D^n$, solved above, as Nature's moves to convert $CG(4)$ to $MCG(4)$. The curtailment rule implies that P1-types' payoff from rejecting P2's stage-III offer is 180. Thus, in the unique $SE(4)$ of $MCG(4)$, all P1-types at stage IV *accept* if $y_2 \geq 180$, irrespective of belief. Therefore, the stage-III $SE(3)$ action is for all P2-types to choose $y_2 = 180$, regardless of belief on P1-types. Therefore, the stage-II $SE(4)$ action is for all P2-types to *accept* if $x_1 \leq (1000 - (600 - 180)) = 580$, regardless of belief on P1-types. Therefore, in $MCG(4)$, at stage-I, 3_1, 4_1, and 5_1 (others replaced by Nature) face an expected payoff of 388 ($= \frac{2\times700}{5} + \frac{3\times180}{5}$) from choosing $x_1 = 700$ versus an expected payoff of 580 from choosing $x_1 = 580$, given that they must have beliefs as per the prior distribution in $SE(4)$. Therefore, in $SE(4)$, at stage-I, 3_1, 4_1, and 5_1 choose $x_1 = 580$. We note SE actions at D^4. This gives us the LFE actions at D^4, which contains the information sets of 3_1, 2_2, 1_2, and 0_1 at stage-I, stage-II, stage-III, and stage-IV, respectively.

Step 5: Fix the LFE actions at $\bigcup_{n=1}^{4} D^n$, solved above, as Nature's moves to convert $CG(5)$ to $MCG(5)$. We now describe the unique $SE(5)$ of $MCG(5)$. All P1-types at stage-V choose $x_3 = 360$ irrespective of belief, because their stage-V payoff is increasing in x_3 according to the curtailment rule. The curtailment rule implies that choosing $x_3 = 360$ gives P1-types a payoff of 180 in the last stage of $MCG(5)$. Therefore, the stage-IV $SE(5)$ action is for all P1-types to *accept* if $y_2 \geq 180$, regardless of belief on P2-types. Given this, the stage-III $SE(5)$ action is for all P2-types to choose $y_2 = 180$. Therefore, the stage-II $SE(5)$ action for all P2-types is to *accept* if $x_1 \leq 580$. Therefore, in $SE(5)$, all P1-types choose $x_1 = 580$ given beliefs determined by the prior distribution on P2-types. We note SE actions at D^5. This gives us the LFE actions at D^5, which contains the information sets of 4_1, 3_2, 2_2, 1_1, and 0_1 at stage-I, stage-II, stage-III, stage-IV, and stage-V, respectively.

Last step: Fix the LFE actions at $\bigcup_{n=1}^{5} D^n$, solved above, as Nature's moves to convert Γ to $MCG(6)$. We now describe the unique $SE(6)$ of $MCG(6)$. Any player-type who observes $MCG(6)$ has no foresight limitation. Thus, in $SE(6)$, all P2-types at stage VI *accept* $x_3 \leq 360$. Given this, the stage-V $SE(6)$ action of all P1-types is $x_3 = 360$, regardless of belief on P2-types. Therefore, the stage-IV $SE(6)$ action of all P1-types is to *accept* $y_2 \geq 360$. Note that 3_2 and 4_2, moving at stage-III, know that if $x_1 = 1000$, then with probability one, P1's type is 0_1, who will accept $y_2 = 180$ in stage-IV. Thus, conditional on $x_1 = 1000$, 3_2 and 4_2 offer $y_2 = 180$. Conditional on $x_1 = 700$, 3_2 and 4_2, moving at stage-III, know that P1's type is 1_1 or 2_1 with probability $\frac{1}{2}$ each. Thus, their stage-III offer of 180 will be rejected by 2_1 with probability $\frac{1}{2}$, and lead to an expected payoff of 210 $(= \frac{(600-180)}{2} + \frac{0}{2})$, which is less than the payoff from proposing $(360, 240)$, and getting a payoff of 240 for sure. Thus, conditional on $x_1 = 700$, 3_2 and 4_2 offer $y_2 = 360$ in stage-III. Therefore in $SE(6)$, in stage-II, 4_2 accepts $x_1 \leq 760$. If 4_2 receives $x_1 \notin \{1000, 700, 580\}$ at stage-II, i.e. off-LFE, 4_2 believes P1's type must be 5_1. In stage-I, 5_1 evaluates the expected payoff from $x_1 = 580$, or 700, or 760, given that his beliefs on P2-types are identical to the prior distribution on P2-types. At stage-I, in $SE(6)$, 5_1 chooses the expected payoff maximizing option: $x_1 = 580$. We note the $SE(6)$ actions at all the information sets not in $\bigcup_{n=1}^{5} D^n$. Steps 1-6 give us the LFE actions for all the information sets of Γ, which completes the LFE strategy profile stated in Table 1.

Proof of Proposition 6.

Lemma 1: *Suppose Γ_0 has the payoff structure P and we replace the payoff profile after "pass" at stage-s with $(x_{s+1},\ y_{s+1})$; where $(x_{s+1},\ y_{s+1})$ is the payoff profile calculated, using the $\frac{min+max}{2}$ curtailment rule. Then, $min\{x_{s+1},\ y_{s+1}\}>max\{(a_i)_{i\leq s},(b_i)_{i\leq s}\}$ holds $\forall\ s < S$.*

Player-type	Stage I	Stage II	Stage III	Stage IV	Stage V	Stage VI
0_1	$x_1 = 1000$			$\underline{Y_2} = 180$	$x_3 = 360$	
1_1	$x_1 = 700$			$\underline{Y_2} = 180$	$x_3 = 360$	
2_1	$x_1 = 700$			$\underline{Y_2} = 360$	$x_3 = 360$	
3_1	$x_1 = 580$			$\underline{Y_2} = 360$	$x_3 = 360$	
4_1	$x_1 = 580$			$\underline{Y_2} = 360$	$x_3 = 360$	
5_1	$x_1 = 580$			$\underline{Y_2} = 360$	$x_3 = 360$	
0_2		$\overline{X_1} = 700$	$y_2 = 0$			$\overline{X_3} = 360$
1_2		$\overline{X_1} = 700$	$y_2 = 180$			$\overline{X_3} = 360$
2_2		$\overline{X_1} = 580$	$y_2 = 180$			$\overline{X_3} = 360$
3_2		$\overline{X_1} = 580$	$y_2 = 180$ if $(x_1 = 1000)$ else $y_2 = 360$			$\overline{X_3} = 360$
4_2		$\overline{X_1} = 760$	$y_2 = 180$ if $(x_1 = 1000)$ else $y_2 = 360$			$\overline{X_3} = 360$

Table 1. LFE Strategies for the Sequential Bargaining Game
Notes: $(x_1, 1000 - x_1)$, and $(x_3, 360 - x_3)$ are the first and third period LFE proposals, respectively, of the first-mover. $\overline{X_1}$ and $\overline{X_3}$ are the maximum first and third period demands, respectively, of the first-mover such that $(\overline{X_1}, 1000-\overline{X_1})$ and $(\overline{X_3}, 360-\overline{X_3})$ are acceptable to the second-mover in those periods in LFE. $(y_2, 600 - y_2)$ is the second-period LFE proposal of the second-mover. $\underline{Y_2}$ is the minimum second-period offer of the second-mover such that $(\underline{Y_2}, 600 - \underline{Y_2})$ is acceptable to the first-mover in the second period in LFE.

Lemma 1 follows straightforwardly due to the properties of payoff structure P.[45] For example, curtailing Figure 7 at stage-3, we get $(x_4,\ y_4) = (132, 72)$, the minimum of which, 72, is higher than the maximum number in {(4,1), (2,8), (16,4)}: 16. Due to Lemma 1, in any Curtailed Game with less than S stages, the highest payoff for both players occurs if both players *pass* in all stages of the Curtailed Game. So, irrespective of ρ (Nature's prior distribution on N), for all $CG(1),\ MCG(2), ..., MCG(S-1)$, there is a unique Sequential Equilibrium: All player-types *pass* with probability 1 at all stages. Thus all limited-foresight types choose *pass* if their total foresight is strictly less than S and they cannot observe Γ. This implies that in $MCG(S)$ (modified version of Γ), according to ρ^S, any Nature's move, at any non-initial node, specifies the pure action: *pass.*

[45]Lemma 1 also holds for several other curtailment rules, including a "mean of stage-wise means" rule followed in an earlier version of this paper. The $\frac{min+max}{2}$ rule is only significant for Lemma 1 in the proof of Proposition 6. Therefore, Proposition 6 holds for several other curtailment rules too.

The decisive information sets of $MCG(S)$ are those where the player-types moving there have total foresight equal to S. We (loosely) refer to the player-types moving at these information sets as rational types. In what follows, we use Proposition 2, and denote the identical beliefs of all rational-types at their corresponding information sets in stage-i as r_i. To show that we can use Proposition 2, we need to show that all the information sets of $MCG(S)$ at stage-i are corresponding information sets. To see this, note that different types of the same player move at stage$-i$, and all the information sets at stage-i are generated from the same action sequence of the underlying Centipede game with perfect information: *pass* played $(i-1)$ times. Thus, by Definition 6, all rational types' information sets at stage-i are corresponding information sets; and by Proposition 2, LFE beliefs of these rational types at corresponding information sets are identical. Let r_i denote the identical belief of every rational player-type at stage-i that at stage-$(i+1)$ the opponent will be *Nature*. By Lemma 1, if *Nature* moves in stage-$(i+1)$ on behalf of a limited-foresight opponent-type and $(i+1) < S$ holds, then *Nature* will choose the pure action *pass* in stage-$(i+1)$. Let p_i denote the identical probability put on the action *pass* by every rational player-type at stage-i. To show Proposition 6, we only need to show that in any SE of $MCG(S)$, it cannot be the case that $p_i = 0$ holds for any $i = 1, ..., (S-3)$. Lemma 2 will be useful in showing this.

Lemma 2: *For any Sequential Equilibrium (σ^S, b^S) of $MCG(S)$: (a) If σ^S implies that $p_i = 1$ holds, then σ^S must imply that $p_s = 1$ holds, for $s \leq i$, where $i = 1, ..., S$. (b) If b^S is such that at stage-i, $r_i > \eta_i$ holds, then σ^S must imply that $p_i = 1$ holds, for $i = 1, ..., S-2$.*

Proof of Lemma 2(a): Let (σ^S, b^S) imply that $p_i = 1$. That is, suppose all rational player-types *pass* with probability 1 at stage-i. Then, sequential rationality of SE strategies implies that according to σ^S, irrespective of beliefs, $p_{i-1} = 1$ must hold. This is because any rational type's choice to *pass* at stage-$(i-1)$ is going to be reciprocated by *pass* with probability 1 at stage-i. Therefore, the payoff from *pass* at stage-$(i-1)$ is at least a_{i+1};

and $a_{i+1} > a_{i-1}$ holds, where a_{i-1} is the payoff from *take* at stage-$(i-1)$. Repeating this argument, $p_{i-1} = 1$ implies that due to sequential rationality of σ^S, we must have $p_{i-2} = 1$, and so on for all $s \leq i$. *Q.E.D.*

Proof of Lemma 2(b): Let $i \in \{1, .., (S-2)\}$. In $MCG(S)$, let v_i be the value to the rational types moving at stage-i given SE play from stage-i on. By sequential rationality of σ^S, it follows that $v_i \geq a_i$ holds for all stages $i \in \{1, .., (S-2)\}$. The payoff from *take* at stage-i is a_i, the expected payoff from *pass* at stage-i is at least $r_i v_{i+2} + [1-r_i]b_{i+1}$ (because, by Lemma 1, *Nature* chooses *pass* with probability 1 at stage-$(i+1)$). So if $r_i > \eta_i$ holds, then:

$$r_i v_{i+2} + [1-r_i]b_{i+1} \geq r_i a_{i+2} + [1-r_i]b_{i+1} > \eta_i a_{i+2} + [1-\eta_i]b_{i+1} = a_i \text{ hold.} \qquad (10)$$

The strict inequality in (10) follows because $a_{i+2} > b_{i+1}$ holds by payoff structure P, and $r_i > \eta_i$ holds. The last equality in (10) follows from the definition of η_i. Note that (10) implies that the expected payoff from *pass* at stage-i, which is greater than the leftmost term of (10), is strictly greater than the payoff from *take* at stage-i: a_i. So, by the sequential rationality of σ^S, $p_i = 1$ must hold. *Q.E.D.*

Given Lemma 1 and Lemma 2, Proposition 6 must also hold for rational player-types. This is because for any SE of $MCG(S)$, (σ^S, b^S), it cannot be the case that σ^S implies $p_i = 0$ at some $i \leq S-3$. We show this by contradiction. Suppose σ^S implies $p_i = 0$ at some $i \leq S-3$. Without loss of generality, suppose player A moves at stage-i. At stage-i, $(S-i)$ A-types have total foresight strictly less than S; and at stage-$(i+2)$, $(S-i-2)$ of those A-types have total foresight strictly less than S.[46] By Lemma 1, all player-types whose total foresight is strictly less than S, choose *pass* with probability 1 at all stages, until their total foresight is S. Thus, if $p_i = 0$, then at stage-$(i+1)$, the player moving there faces only $(S-i)$

[46] $\{0_A, ..., (S-i-1)_A\}$ all have total foresight strictly less than S at stage-i. $\{0_A, ..., (S-i-3)_A\}$ all have total foresight strictly less than S at stage-$(i+2)$.

equiprobable limited-foresight A-types who chose *pass*, out of whom $(S-i-2)$ A-types will choose *pass* again in stage-$(i+2)$. So, by the consistency of SE beliefs, if $p_i = 0$, then

$$r_{i+1} = \frac{S-i-2}{S-i} \geq \frac{1}{3} > \eta_{i+1} = \frac{1}{7} \text{ holds.}$$

So, by Lemma 2(b), we must have $p_{i+1} = 1$ according to σ^S. But then Lemma 2(a) implies $p_i = 1$ must hold according to σ^S, which is a contradiction. So in any SE of $MCG(S)$, and thus in any LFE, it cannot be the case that $p_i = 0$ holds for $i = 1, ..., (S-3)$. *Q.E.D.*