Management Science

A Dynamic Level-k Model in Sequential Games

Teck-Hua Ho, Xuanming Su,

Please scroll down for article—it is on subsequent pages

A Dynamic Level-*k* Model in Sequential Games

Teck-Hua Ho
National University of Singapore, Singapore 119077; and University of California, Berkeley, Berkeley, California 94720,
hoteck@haas.berkeley.edu

Xuanming Su
The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104,
xuanming@wharton.upenn.edu

Backward induction is a widely accepted principle for predicting behavior in sequential games. In the classic example of the "centipede game," however, players frequently violate this principle. An alternative is a "dynamic level-*k*" model, where players choose a rule from a rule hierarchy. The rule hierarchy is iteratively defined such that the level-*k* rule is a best response to the level-$(k-1)$ rule, and the level-∞ rule corresponds to backward induction. Players choose rules based on their best guesses of others' rules and use historical plays to improve their guesses. The model captures two systematic violations of backward induction in centipede games, limited induction and repetition unraveling. Because the dynamic level-*k* model always converges to backward induction over repetition, the former can be considered to be a tracing procedure for the latter. We also examine the generalizability of the dynamic level-*k* model by applying it to explain systematic violations of backward induction in sequential bargaining games. We show that the same model is capable of capturing these violations in two separate bargaining experiments.

Key words: level-*k* models; learning; sequential games; backward induction; behavioral game theory
History: Received March 15, 2011; accepted August 3, 2012, by Peter Wakker, decision analysis. Published online in *Articles in Advance* November 9, 2012.

1. Introduction

In many settings, players interact with one another over multiple stages. Researchers often model these settings as sequential games and invoke the principle of backward induction to predict behavior at each stage of these games. Under backward induction, players reason backward, replace each subgame of a sequential game by its optimal payoff, always choose optimally within each subgame, and use this iterative process to determine a sequence of optimal actions. Each player follows this procedure betting on others doing the same. This "divide and conquer" algorithm simplifies the game analysis and generates a sharp prediction of game play for any sequential game.

However, subjects who are motivated by substantial economic incentives often violate backward induction even in simple sequential games. One such example is the "centipede" game (Rosenthal 1981) (see the top panel of Figure 1). In this game, there are two players (A and B) and four decision stages. Players are endowed with an initial pot of $5. In Stage I, player A has the property rights to the pot. She can choose either to end the game by taking 80% of the pot (and leaving the remaining 20% to player B) or to allow the pot to double by passing the property rights to player B. In Stage II, it is now player B's turn to make a similar decision. Player B must now

decide whether to end the game by taking 80% of $10 or to let the pot double again by passing the property rights back to player A. This social exchange process leads to large financial gains as long as both players surrender their property rights at each stage. At Stage IV, player B can either take 80% of $40 (i.e., $32) or pass and be left with 20% of a pot of $80 (i.e., $16).

Backward induction generates a clear prediction for this game by starting the analysis in the last stage. Player B should take at Stage IV because 80% of $40 (i.e., $32) is larger than 20% of $80 (i.e., $16). Anticipating this choice and replacing this subgame with the corresponding payoff vector, player A should take at Stage III because 80% of $20 (i.e., $16) is larger than 20% of $40 (i.e., $8). Continuing with this line of logic, backward induction makes a surprising prediction: player A always takes immediately in Stage I and outcome 4 occurs with probability 1 (i.e., outcomes 0–3 will not occur). Moreover, the same prediction holds even if the game continues for more stages and with more dramatic financial gains. For example, the bottom panel of Figure 1 shows the same game with six stages, and the same sharp prediction holds (i.e., outcome 6 occurs with probability 1). Introspection suggests that this prediction is unlikely to occur. This is so because as long as the game proceeds to Stage III, both players would have earned more money than the backward induction outcome.

Figure 1   Four-Stage (Top) and Six-Stage (Bottom) Centipede Games
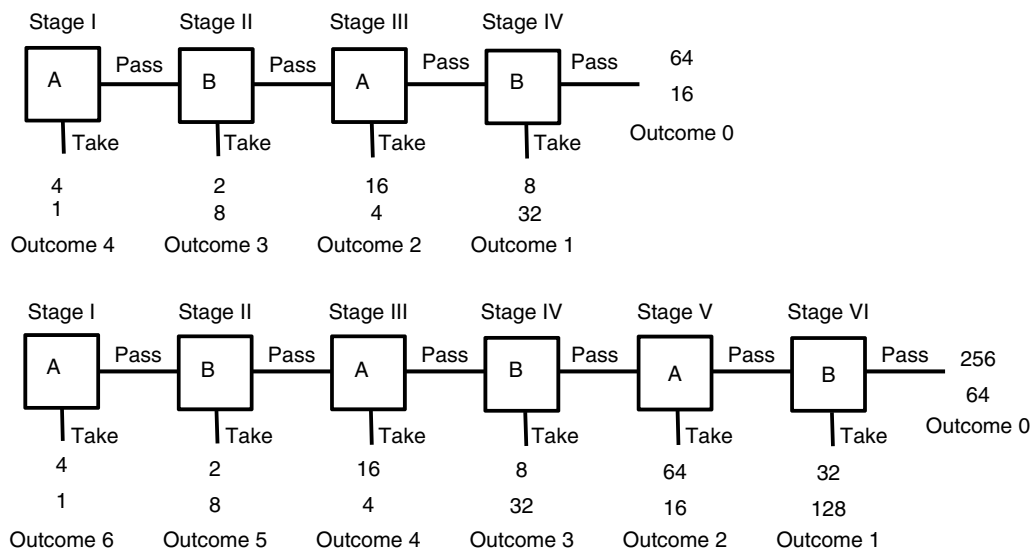
Indeed, very few subjects (about 6%) chose to take immediately in experimental centipede games conducted by McKelvey and Palfrey (1992). Many subjects instead chose to take in the intermediate stages, approximately halfway through the games (i.e., leading to outcomes 2 and 3 in four-stage games and outcomes 3 and 4 in six-stage games). Clearly, this pattern of behavior runs counter to backward induction. Moreover, the observed behavior frequently led to higher cash earnings for all subjects. In contrast, those who obeyed backward induction by taking immediately received a substantially lower payoff.

There are two stylized facts concerning the violations of backward induction. First, players violate backward induction less in a game with fewer subgames (or stages); that is, players' behaviors deviate less from backward induction in simpler games. For instance, we observe fewer violations of backward induction in four-stage than in six-stage games (see Figure 1). We call this behavioral tendency *limited induction*. Second, players unravel as they play the same game repeatedly over multiple rounds; that is, players' behaviors converge toward backward induction over repetition. For instance, we observe fewer violations of backward induction in the last round than in the first round of the experimental centipede games in Figure 1. This behavioral tendency is termed *repetition unraveling*. The inability of backward induction to account for these two empirical stylized facts poses significant modeling challenges.

This paper proposes an alternative to backward induction, a "dynamic level-k" model, that generalizes backward induction and accounts for limited induction and repetition unraveling in centipede games. In the dynamic level-k model, players choose a level-k rule, L_k ($k = 0, 1, 2, 3, \ldots,$), from a set of iteratively defined rules. Each rule prescribes an action

at each subgame. The rule hierarchy is defined such that the level-k rule best responds to the level-$(k-1)$ rule, and the level-∞ rule corresponds to backward induction.[1] This iterative definition of a rule hierarchy is usually applied to one-shot games (e.g., 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), but we extend the approach to sequential games. Players choose a rule based on their beliefs of others' rules, so they essentially are subjective expected utility maximizers.

Players are heterogeneous in that they have different initial guesses of others' rules and consequently choose different initial rules. The distribution of the initial guesses is assumed to follow any arbitrary discrete distribution. These initial guesses are updated according to Bayes' rule based on game history. Consequently, players develop more accurate guess of others' rules and may choose different rules over time. In this way, unlike static level-k models, our model is made dynamic by incorporating elements of belief-based learning models.

We prove that the dynamic level-k model can account for limited induction and repetition unraveling properties in centipede games. Consequently, our model can explain why subjects choose to pass in earlier stages in these games. Such behavior is considered paradoxical under backward induction but is consistent with the dynamic level-k model. In addition, the

---

[1] Whereas Nagel (1995), Ho et al. (1998), Costa-Gomes et al. (2001), Costa-Gomes and Crawford (2006), and Crawford and Iriberri (2007a, b) assume that the level-k rule only best responds to the level-$(k-1)$ rule, Stahl and Wilson (1995), Stahl (1996), and Camerer et al. (2004) assume that level-k rule best responds to all lower-level rules. This paper adopts the former approach.

dynamic level-*k* model is able to capture the empirical stylized fact that behavior will eventually converge to backward induction over repetition, and hence the former can be considered as a tracing procedure for the latter.

We also fit our model using experimental data on centipede games from McKelvey and Palfrey (1992) and find that our model fits the data significantly better than backward induction and the static level-*k* model, where no dynamics is allowed. In addition, we rule out two alternative explanations, including the reputation-based model of Kreps et al. (1982) and a model allowing for inequity aversion (Fehr and Schmidt 1999). Overall, it appears that the dynamic level-*k* model can be an empirical alternative to backward induction for predicting behaviors in centipede games.

To investigate the generalizability of the dynamic level-*k* model to other sequential games, we investigate whether the same dynamic level-*k* model can be used to explain violations of backward induction in sequential bargaining games (Stahl 1972, Rubinstein 1982). We find that the same model can explain both the initial offers and the shift in offers over time in two separate experiments. Hence, the dynamic level-*k* model has applicability beyond the centipede games.

The rest of this paper is organized as follows. Section 2 discusses the backward induction principle and its violations. Section 3 formulates the dynamic level-*k* model and applies it to explain paradoxical behaviors in centipede games. Section 4 fits the dynamic level-*k* model to data from experimental centipede games and rules out two alternative explanations. Section 5 applies the dynamic level-*k* model to explain violations of backward induction in sequential bargaining games. Section 6 concludes.

2. Violations of Backward Induction

Backward induction uses an iterative process to determine an optimal action at each subgame of a sequential game. The predictive success of this iterative reasoning process hinges on players' complete confidence in others applying the same logic in arriving at the backward induction outcomes (Aumann 1995). If players have doubts about others applying this same reasoning process, it may be in their best interest to deviate from the prescription of backward induction. Indeed, subjects do, and profitably so in many experiments.

If player i chooses a behavioral rule L^i that is different from backward induction (L_∞), one would like to develop a formal measure to quantify this deviation. Consider a centipede game G_S with S subgames. We can define the deviation for a set of behavioral rules L^i ($i = A, B$), one for each player, in centipede game G_S as

$$\delta(L^A, L^B, G_S) = \frac{1}{S} \sum_{s=1}^{S} D_s(L^i, L_\infty), \qquad (1)$$

where $D_s(L^i, L_\infty)$ is 1 if player i chooses an action at subgame s that is different from the prescription of backward induction and 0 otherwise. Note that the measure varies from 0 to 1, where 0 indicates that players' actions perfectly match the predictions of backward induction, and 1 indicates that none of the players' actions agree with the predictions of backward induction.

Let us illustrate the deviation measure using a four-stage centipede game. Let the behavioral rules adopted by player A and B be $L^A = \{P, -, T, -\}$ and $L^B = \{-, P, -, T\}$, respectively (that is, player A will pass in Stage I and take in Stage III, and player B will pass in Stage II and take in Stage IV). Then the game will end in Stage III (i.e., outcome 2). The deviation will be $\delta(L^A, L^B, G) = (1/4)[1 + 1 + 0 + 0] = 1/2$. Similarly, if $L^A = \{P, -, T, -\}$ and $L^B = \{-, T, -, T\}$, then the game will end in Stage II (i.e., outcome 3). This gives $\delta(L^A, L^B, G) = (1/4)[1 + 0 + 0 + 0] = 1/4$, which is smaller. Note that the latter behavioral rules are closer to backward induction than the former behavioral rules.

Using the above deviation measure, we can formally state the two systematic violations of backward induction as follows:

1. *Limited induction*: Consider two centipede games G and G', where G' is a proper subgame of G. The deviation from backward induction is equal or larger in G than in G'; that is, the deviation from backward induction increases in the number of stages or subgames S. Formally, for two players who adopt the same set of behavioral rules ($L^i, i = A, B$) in games G and G', we have $\delta(L^A, L^B, G) \geq \delta(L^A, L^B, G')$. Consequently, a good model must predict a larger deviation in G than in G' to be behaviorally plausible.

2. *Repetition unraveling*: If a game G is played repeatedly, the deviation from backward induction at the tth round converges to zero as $t \to \infty$; that is, repetition unraveling implies $\delta(L^A(t), L^B(t), G) \to 0$ as $t \to \infty$. Therefore, game outcomes will eventually be consistent with backward induction after sufficiently many repetitions.

Let us illustrate limited induction and repetition unraveling using data from McKelvey and Palfrey (1992). These authors conducted an experiment to study behavior in four-stage and six-stage centipede games. Each subject was assigned to one of these games and played the same game in the same role 9 or 10 times. For each observed outcome in a game play, we can compute the deviation from backward

Figure 2   Deviations from Backward Induction in the Four-Stage and Six-Stage Centipede Games

Source. Data from McKelvey and Palfrey (1992).

Figure 3   Deviations from Backward Induction in the First and Last Round of the Four-Stage Centipede Game
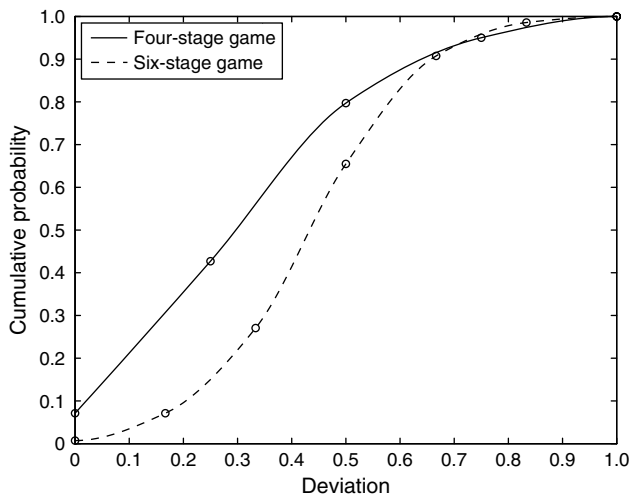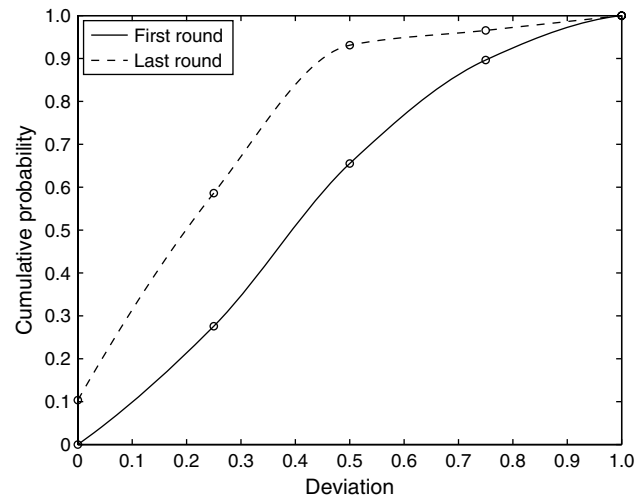

Source. Data from McKelvey and Palfrey (1992).

induction using Equation (1). Because subjects did not indicate what they would have chosen in every stage (i.e., data were not collected using the strategy method), we did not observe what the subjects would have done in subsequent stages if the game ended in an earlier stage. In computing the deviation, we assume that subjects always choose to take in stages beyond where the game ends. Therefore, our measure is a conservative estimate of the deviation from backward induction.

Figure 2 plots the cumulative distributions of deviations from backward induction in the four-stage and six-stage games respectively. The solid line corresponds to the four-stage game, and the dashed line corresponds to the six-stage game. The curve for the six-stage game generally lies to the right of the curve for the four-stage game except for high deviation values. A Kolmogorov–Smirnov test shows that there is a statistically significant difference between the distributions of deviations in the two games. These results suggest that the limited induction property holds in this data set.

Figure 3 plots the cumulative distributions of deviations from backward induction in the first and the final round of the four-stage game. The solid line corresponds to the first round, and the dashed line corresponds to the final round of game plays (similar results occur for the six-stage game). As shown, the curve for the first round lies to the right of the curve for the final round. A Kolmogorov–Smirnov test shows that there is a statistically significant difference between the distributions of deviations in the first and last rounds. These results suggest that the deviation from backward induction decreases over time.

We shall use the deviation measure to establish the main theoretical results below. Specifically, we shall show that the deviation under our dynamic level-k model is smaller in simpler games (i.e., the limited induction property holds) and converges to zero over repetition (i.e., the repetition unraveling property holds).

3. Dynamic Level-k Model

3.1. Model Setup

3.1.1. Rule Hierarchy. We consider a centipede game that has S subgames. Players are indexed by i ($i = A, B$), and subgames by s ($s = 1, \ldots, S$). Players are assumed to adopt a rule that prescribes an action at each subgame s. For example, in the centipede game studied by McKelvey and Palfrey (1992), S is either 4 or 6 and a rule player A adopts in a four-stage game can be $L^A = \{P, -, T, -\}$, which specifies that the player will pass in Stage I and take in Stage III. Players choose a rule from a rule hierarchy. Rules are denoted by L_k ($k = 0, 1, 2, 3, \ldots,$). The L_0 rule prescribes naive or uniform randomization among all available actions in every subgame s and all other higher-level rules, $L_1, L_2, \ldots,$ are generated from iterative best responses. Specifically, the L_k ($k \geq 1$) rule is a best response to the L_{k-1} rule *at every subgame*, including those subgames that may never be reached;[2] that is, to determine a player's decisions under rule L_k, we first assume that the other

---

[2] Note that level-k can best respond to level-$(k-1)$ and yet "pass" at later nodes because once the former chooses to "take" at a node the latter can choose whatever action at later nodes because those nodes will never be reached. To keep the analysis simple, we impose the plausible assumption that level-k will best respond to

Table 1   Rule Hierarchy in Four-Stage Centipede Games

Level rule	Player A	Player B
L_0	$\{R, -, R, -\}$	$\{-, R, -, R\}$
L_1	$\{P, -, P, -\}$	$\{-, P, -, T\}$
L_2	$\{P, -, T, -\}$	$\{-, P, -, T\}$
L_3	$\{P, -, T, -\}$	$\{-, T, -, T\}$
L_4	$\{T, -, T, -\}$	$\{-, T, -, T\}$

Notes. R, randomize; *P*, pass; *T*, take.

Table 2   Rule Hierarchy in Six-Stage Centipede Games

Level rule	Player A	Player B
L_0	$\{R, -, R, -, R, -\}$	$\{-, R, -, R, -, R\}$
L_1	$\{P, -, P, -, P, -\}$	$\{-, P, -, P, -, T\}$
L_2	$\{P, -, P, -, T, -\}$	$\{-, P, -, P, -, T\}$
L_3	$\{P, -, P, -, T, -\}$	$\{-, P, -, T, -, T\}$
L_4	$\{P, -, T, -, T, -\}$	$\{-, P, -, T, -, T\}$
L_5	$\{P, -, T, -, T, -\}$	$\{-, T, -, T, -, T\}$
L_6	$\{T, -, T, -, T, -\}$	$\{-, T, -, T, -, T\}$

Notes. R, randomize; *P*, pass; *T*, take.

player follows rule L_{k-1} and then solve the single-player dynamic program that results. In the limit, L_∞ corresponds to backward induction. Tables 1 and 2 show the respective rule hierarchy in four-stage and six-stage centipede games. In Table 1, for example, L_4 for player A is $\{T, -, T, -\}$. Similarly, in Table 2, L_6 for player A is $\{T, -, T, -, T, -\}$. Note that in both cases each rule requires that player A takes in all subgames after Stage I (i.e., "take" in Stage III in four-stage centipede games, and "take" in Stages III and IV in six-stage centiepde games) even though those nodes will never be reached.

Under the dynamic level-k rule hierarchy, the deviation from backward induction (see Equation (1)) is smaller if player i adopts a higher-level rule (while others keep their rule at the same level). In fact, ceteris paribus, the deviation from backward induction is (weakly) monotonically decreasing in k. In other words, the level of a player's rule captures its closeness to backward induction. Note that L_k prescribes the same behavior as backward induction in any centipede game with k or fewer subgames. In this regard, L_k can be viewed as a limited backward induction rule that only works for simpler centipede games.

If it is common knowledge that everyone employs backward induction, all players will indeed choose L_∞. However, if players have doubts about others' use of L_∞, it may not be in their best interest to apply the same rule. As a result, it is natural

for players to form beliefs[3] about which less sophisticated rules others will adopt and then determine a best response to maximize their expected payoffs.[4] Put differently, players are subjective expected utility maximizers in the dynamic level-k models.

3.1.2.   Belief Updating. In a typical laboratory experiment, players frequently play the same game repeatedly. After each repetition, players observe the rules used by their opponents and update their beliefs by tracking the frequencies of rules played by opponents in the past. Let player i's rule counts at the end of the tth round be $N^i(t) = (N_0^i(t), \dots, N_S^i(t))$, where $N_k^i(t)$ is the cumulative count of rule L_k that has been used by opponents at the end of round t (Camerer and Ho 1999, Ho et al. 2007). Note that for a centipede game with S stages or subgames, all rules of level S or higher will prescribe the same action at each subgame, and hence we pool them together and collectively call them L_S. Given these rule counts, player i forms a belief $B^i(t) = (B_0^i(t), \dots, B_S^i(t))$, where

$$B_k^i(t) = \frac{N_k^i(t)}{\sum_{k'=0}^{S} N_{k'}^i(t)};$$

$B_k^i(t)$ is player i's belief of the probability that her opponent will play L_k in round $t+1$. The updating equation of the cumulative count at the end of round t is given by

$$N_k^i(t) = N_k^i(t-1) + I(k, t) \cdot 1, \quad \forall k, \qquad (2)$$

where $I(k, t) = 1$ if player i's opponent adopts rule L_k in round t and 0 otherwise. Therefore, players update their beliefs based on the history of game plays. Note that this belief updating process is similar to the standard fictitious play model (Brown 1951). Furthermore, the updating process is consistent with

---

[3] There is a debate about whether players actually form beliefs and best respond to them, or whether they simply exhibit behavior "as if" they were forming beliefs. Although we adopt the former view (because it is so central to game theoretic reasoning), there appears to be some evidence to support the latter. For example, in experimental games, Costa-Gomes and Weizsacker (2008) found that subjects often fail to best respond to their own beliefs, and Charness and Levin (2009) found that subjects are unable to update beliefs "correctly," i.e., perform Bayesian updating.

[4] Subjects' beliefs may depend on their knowledge of their opponents' level of sophistication. Players who play against opponents who are known to be sophisticated will adopt a higher-level rule. For example, Palacios-Huerta and Volij (2009) showed that many chess players in centipede games choose to pass when playing against student subjects, but choose to take immediately when playing against other equally sophisticated chess players. Levitt et al. (2011), however, found the reverse result: chess players choose to pass when they play against each other. This empirical discrepancy could be due to the difference in players' perception of their opponents' level of sophistication.

---

level-$(k-1)$ in all subgames of the subgame whose initial node is where level-$(k-1)$ first chooses "take." We thank a knowledgeable reviewer for suggesting to us to make this assumption explicit.

Bayesian updating involving a multinomial distribution with a Dirichlet prior (Fudenberg and Levine 1998, Camerer and Ho 1999). As a consequence of the updating process, players may adopt different best-response rules over game repetitions t.[5]

Player i chooses the optimal rule L_{k^*} in round $t+1$ from the rule hierarchy $\{L_0, L_1, \ldots, L_S\}$ based on belief $B^i(t)$ to maximize expected payoffs. Let a_{ks} be the specified action of rule L_k at subgame s. Player i believes that action $a_{k's}$ will be chosen with probability $B^i_{k'}(t)$ by the opponent. Hence, the optimal rule chosen by player i is

$$k^* = \arg \max_{k=1,\ldots,S} \sum_{s=1}^{S} \left\{ \sum_{k'=1}^{S} B^i_{k'}(t) \cdot \pi^i_s(a_{ks}, a_{k's}) \right\}, \quad (3)$$

where $\pi^i_s(a_{ks}, a_{k's})$ is player i's payoff at subgame s if i chooses rule L_k and the opponent chooses rule $L_{k'}$ (cf. Camerer et al. 2004).

Note that we model learning across repetitions but not across stages within a game. This is clearly an approximation. For a more general model, a player could potentially update her belief about the opponent' rule across stages as the game unfolds. In the centipede game, player B who expects player A to take immediately will be surprised if the latter passes. If we incorporate within-round learning, this will lead player B to put more weights on the lower-level rules as the game progresses. Nevertheless, we believe that our simpler model is a good starting point for two reasons:

1. A player learns when she is surprised by the opponent's action. This can occur in two ways. First,

a player is surprised when an opponent takes earlier than expected. In this case, the game ends immediately. The above updating process in Equations (2) and (3) already captures this kind of between-round learning. Second, a player is surprised when an opponent unexpectedly passes. However this kind of within-round learning can only happen to player B who expects her opponent to take in Stage I but she passes instead. This is rare because the initial pass in Stage I will only be considered "unexpected" by very high-level-rule players (i.e., level 5 or higher in the four-stage game and level 7 or higher in the six-stage game in round 1). The estimated fraction of players who are level 5 or higher is less than 8%, and the estimated fraction of players who are level 7 or higher is less than 2%.

2. Within-round learning as described above frequently generates prediction that is inconsistent with the observed behavior. Specifically, within-round learning tends to make players more likely to pass after observing a surprising pass by the opponent. This pattern of behavior runs counter to the observed data.

3.1.3. Initial Beliefs.
We need to determine player i's initial belief $B^i(0)$. We define $N^i(0)$ such that $N^i_k(0) = \beta$ for some k and 0 otherwise, where the parameter β captures the strength of the initial belief. In other words, player i places all the initial weight β on a particular rule L_k and zero weight on all other rules. Different players have different guesses about others' level of sophistication and hence place the initial weight on a different k. To capture the heterogeneity in players' initial beliefs of others' rules, the dynamic level-k model allows for any arbitrary discrete distribution. Let $\phi(k)$ denote the proportion of players who hold an initial belief that the opponents' rule is level-k, for $k = 0, \ldots, S$. For example, a $\phi(0)$ proportion of players initially believe that their opponents will play L_0 and thus choose L_1 in round 1. Similarly, a $\phi(k)$ proportion of players initially believe that their opponents will play L_k and thus best respond with L_{k+1}.

We should stress that $\phi(k)$ has a different interpretation here, compared to other level-thinking models. For example, $\phi(0)$ represents the proportion of L_0 players in the static level-k and cognitive hierarchy models, but it represents, in our model, the proportion of players who *believe* that their opponents play L_0 and thus choose L_1 themselves. Unlike the static level-k and cognitive hierarchy models, which capture heterogeneity in players' rules, the dynamic level-k model captures heterogeneity in players' beliefs of others' rules and allows players to always best-respond to their beliefs. Players can and do change their rules as a result of changes in their beliefs of others' rules in the dynamic level-k model.

---

[5] The above updating rule assumes that subjects observe rules chosen by opponents. This is possible if the strategy method is used to elicit subjects' contingent action at each subgame. When the opponents' chosen rules are not observed, the updating process is still a good approximation because subjects may have a good guess of their opponents' chosen rules in most simple games (e.g., centipede games). More generally, the updating of $N^i_k(t)$ depends on whether player i adopts a higher- or lower-level rule than her opponent. If the opponent uses a higher-level rule (e.g., the opponent takes before the player in the centipede game), then we have, similar to the above,

$$N^i_k(t) = N^i_k(t-1) + I(k, t) \cdot 1,$$

where $I(k, t) = 1$ if opponent adopts an action that is consistent with L_k in round t and 0 otherwise. If player i adopts a higher-level rule k^* (e.g., takes before the opponent in the centipede game), the player can only infer that the opponent has chosen some rule that is below k^*. Then we have

$$N^i_k(t) = N^i_k(t-1) + I(k \le k^*) \cdot \frac{N^i_k(t-1)}{\sum_{k'=0}^{k^*} N^i_{k'}(t-1)},$$

where $I(k \le k^*) = 1$ if $k \le k^*$ and 0 otherwise. This updating process assigns a belief weight to all lower-level rules that are consistent with the observed outcome. The weight assigned to each consistent rule is proportional to its prior belief weight. For this alternative updating process, the main results (i.e., Theorems 1 and 2) continue to hold. We use the simpler updating process in our empirical estimation.

Because all players best respond given their beliefs, L_0 will not be chosen by any player and only occurs in the minds of the higher-level players. Using a general discrete distribution, Costa-Gomes and Crawford (2006) and Crawford and Iriberri (2007a, b) show that the estimated proportion of players who adopt the L_0 rule is frequently zero. In agreement with this finding, our proposed dynamic level-k model assumes that there are no L_0 players. In this way, the dynamic level-k model is both an empirically validated and theoretically justified model of strategic behavior.

3.1.4. Comparison with Other Models. The dynamic level-k model is different from the static level-k and cognitive hierarchy models in three fundamental aspects. First, players in our model are not endowed with a specific thinking type; that is, players in our model are cognitively capable of choosing any rule and always choose the rule level that maximizes their expected payoff. In other words, players in our model are not constrained by reasoning ability. Second, players in our model may be aware of others who adopt higher-level rules than themselves. In other words, a player who chooses L_k may recognize that there are others who choose L_{k+1} or higher but still prefers to choose L_k because there is a large majority of players who are L_{k-1} or below. On the other hand, the static level-k and cognitive hierarchy (CH) models assume that players always believe they are the highest-level thinkers (i.e., the opponents are always of a lower-level rule). Third, unlike the static level-k and CH models, players in the dynamic level-k model may change their rules as they collect more information and update beliefs about others. Specifically, a player who interacts with opponents of higher-level rules may advance to a higher rule. Similarly, a players who interacts with opponents of lower-level rules may switch to a lower-level rule to maximize their expected payoffs.

3.1.5. Summary. In summary, the dynamic level-k model is characterized by the parameter β and the distribution $\phi(\cdot)$. The parameter β is the strength of the initial belief, which determines players' sensitivity to game history. A higher β implies a lower level of sensitivity to game history. The distribution $\phi(\cdot)$ captures the degree of heterogeneity in players' initial beliefs. The distribution $\phi(\cdot)$ is likely to depend on the sophistication of the player population. The dynamic level-k model nests several well-known special cases. When $\phi(\infty) = 1$, the model reduces to backward induction. If $\beta = \infty$, players have a stubborn prior and never respond to game history. This reduces our model to a variant of the static level-k model. Consequently, we can empirically test whether these special cases are good approximations

of behavior using the standard generalized likelihood principle.

3.2. Theoretical Implications

We now apply the dynamic level-k model to explain violations of backward induction in centipede games. McKelvey and Palfrey (1992) studied centipede games with an even number of stages (e.g., four and six stages). Hence, we focus on centipede games that have an even number of stages. In these games, under the L_0 rule, players randomly choose between passing and taking with equal probabilities at every stage. A player who believes her opponent uses L_0 will maximize her payoff by adopting $L_1 = \{P_1, P_2, \ldots, P_{S-1}, T_S\}$. (Note that odd-numbered components apply to player A, and even-numbered components apply to player B. See also Tables 1 and 2.) By definition, L_k best responds to L_{k-1} so that $L_k = \{P_1, \ldots, P_{S-k}, T_{S-k+1}, \ldots, T_S\}$. Therefore, in a centipede game with S stages, $L_S = L_{S+1} = \cdots = L_\infty$. Put differently, all rules L_S or higher prescribe the same action at each stage as L_∞ (the backward induction rule). Consequently, we pool all these higher-level rules together and collectively call them L_S.

Players choose rules from the rule hierarchy L_k ($k = 0, 1, 2, 3, \ldots,$). Let $L^i(t)$ be the rule of player i (where $i = A, B$) in round t. Then $\delta(L^A(1), L^B(1), G)$ is the deviation from backward induction L_∞ in game G in the first round. For example, if players A and B both play L_2 in the first round, the deviation is $1/2$ in a four-stage and $2/3$ in a six-stage centipede game; that is, we have $\delta(L_2^A(1), L_2^B(1), G_6) = 2/3 > \delta(L_2^A(1), L_2^A(1), G_4) = 1/2$, where G_4 and G_6 denote the four-stage and six-stage games, respectively. In general, we have the following theorem:

THEOREM 1. *In centipede games, the dynamic level-k model implies that the limited induction property is always satisfied. Formally, if each player holds the same initial beliefs in both games G_S and $G_{S'}$ with $S < S'$, then the expected deviation from backward induction in G_S, $\delta(L^A(1), L^B(1), G_S)$, is always smaller than in $G_{S'}$, $\delta(L^A(1), L^B(1), G_{S'})$.*

PROOF. See Appendix A.

Theorem 1 suggests that the dynamic level-k model gives rise to a smaller deviation from backward induction in a centipede game with a smaller number of stages. This result is consistent with the data presented in Figure 2. Appendix A gives the detailed proof but the basic idea of the proof is outlined here. Given any combination of rule levels $L^A(1)$ and $L^B(1)$ for the players, let $K^A(1) = 2 \cdot \lfloor L^A(1)/2 \rfloor$ and $K^B(1) = 2 \cdot \lceil L^B(1)/2 \rceil - 1$. The outcome is identical in both games (counting from the last stage) (see Figure 1). Specifically, the game outcome is $z = \max\{K^A(1), K^B(1)\}$, and

the number of actions that are inconsistent with backward induction is $S - z$. As a consequence, G_6 has a larger deviation than G_4 (see Equation (1)). Because the initial distribution of beliefs is the same in both games (i.e., having the same $\phi(\cdot)$), the initial expected deviation must be higher in G_6.

Figure 3 shows that the expected deviation from backward induction becomes smaller over time. There is a question of whether this trend will persist and eventually unravel to backward induction outcome. The following theorem formally states that this is indeed the case for the dynamic level-k model.

THEOREM 2. *For a centipede game with S stages, if β is finite, the dynamic level-k model implies that the repetition unraveling property is always satisfied. Formally, the deviation from backward induction $\delta(L^A(t), L^B(t), G)$ converges to zero and all players will choose $L_S = L_\infty$ as $t \to \infty$; that is, players will eventually take in every stage.*

PROOF. See Appendix A.

Theorem 2 states that the dynamic level-k model satisfies the repetition unraveling property in centipede games.[6] The basic idea of the proof is outlined here. No players choose L_0. As a consequence, L_1 players will learn that other players are L_1 or higher. Using our notation, this means that $B_0^i(t)$ will decline over time and the speed of decline depends on the initial belief weight β. For a specific β, there is a corresponding number of rounds after which all L_1 players will move up to L_2 or higher. No players will then choose L_0 and L_1. In the same way, L_2 players will learn that other players are L_2 or higher and will eventually move to L_3 or higher. Consequently, we will see a "domino" phenomenon whereby lower-level players will successively disappear from the population. In this regard, players believe that others become more sophisticated over time and correspondingly do

---

[6] Fey et al. (1996) studied a constant-sum version of the centipede game in which the total payoff to the two players remains constant over stages. The players are initially endowed with two equal piles of cash. Similar to the regular centipede game, each player takes turn to either take or pass. Each time a player passes, one-fourth of the smaller pile is transferred to the larger pile. When a players takes, she receives the larger pile of the cash, and the game ends. The backward induction principle predicts that players should take immediately in the first stage in this game. However, subjects did not take immediately, but unraveled toward the backward induction outcome over time. Applying the dynamic level-k model to this centipede game, we obtain the following rule hierarchy: L_0 prescribes random choice, L_1 prescribes passing in the first stage and taking in every other stage, and L_2 and above corresponds to backward induction (i.e., taking in every stage). Given this rule hierarchy, the same reasoning as in Theorem 2 can be used to show that the repetition unraveling property holds. This suggests that the dynamic level-k model can also be used to explain the learning pattern in constant-sum centipede games. We thank a reviewer for pointing this out.

so themselves. In the limit, all players converge to L_S (and the learning process ceases).

The proof also reveals an interesting insight. The number of repetitions it takes for L_k to disappear is increasing in k. For example, when $\beta = 1$, it takes 12 repetitions for L_1 to disappear and another 50 repetitions for L_2 to disappear. Each higher-level rule takes an exponentially longer time to be eliminated from the population.[7] This result suggests that repetition unraveling occurs rather slowly in centipede games. In the experiments conducted by McKelvey and Palfrey (1992), there are only 10 game rounds, and hence we observe only a slight convergence toward backward induction. In §4, we fit the dynamic level-k model to yield $\beta < 1$ (0.16 for Caltech subjects and 0.80 for Pasadena Community College (PCC) subjects), and we find that only the L_1 rule disappears in their data set.

Theorem 2 above shows that all players will choose $L_S = L_\infty$ as $t \to \infty$. However, during the transient phase, it is possible for players to switch from L_k to $L_{k'}$ ($k' < k$) if they repeatedly encounter opponents of lower-level rules. This phenomenon is occasionally observed in the data and can be accommodated by our model.

We have proven the limited induction and repetition unraveling properties of the dynamic level-k model by adopting a commonly used specification for L_0 that prescribes uniform randomization. This same specification is used in prior level-thinking and cognitive hierarchy models (e.g., Ho et al. 2004, Costa-Gomes and Crawford 2006, Crawford and Iriberri 2007b). From a modeling standpoint, this level-0 rule represents a natural starting point from which we can iteratively derive more sophisticated rules. However, it is worthwhile to ask whether the above results are robust to alternative specifications of L_0. For centipede games, a possible alternative is pure altruism behavior, i.e., L_0 players always pass in every stage. In fact, McKelvey and Palfrey (1992) assumed there is a proportion of such altruistic players to reconcile deviations of their data from standard equilibrium models. To investigate the robustness of our results, we investigate an alternative rule hierarchy generated from L_0 exhibiting pure altruism (i.e., $L_0 = \{P_1, P_2, \ldots, P_{S-1}, P_S\}$). Interestingly, we find that this alternative L_0 generates an identical rule hierarchy L_k ($k \geq 1$); that is, for $k \geq 1$, $L_k = \{P_1, \ldots, P_{S-k}, T_{S-k+1}, \ldots, T_S\}$ as obtained above. Recall that in the dynamic level-k model, although L_0 may be part of players' beliefs, only rules L_1 and higher

---

[7] In our model, the cumulative rule counts $N_k^i(t)$ do not decay over time. If decay is allowed, that is, $N_k^i(t) = \delta \cdot N_k^i(t-1) + I(k, t) \cdot 1$, then unraveling can occur at a much faster rate. For example, if $\delta = 0$, it may take only one repetition for each successively lower-level rule to disappear.

will actually be chosen by players. The identical rule hierarchy, however, does not imply that the dynamics of learning remain unchanged. In fact, we find that this alternative specification of L_0 leads to a different rate of learning for higher-level rules. Despite this difference, we can still prove that Theorems 1 and 2 discussed above continue to hold. Furthermore, the qualitative nature of our empirical results reported below remain unchanged under this alternative specification.

4. An Empirical Application to Centipede Games

We use the data from experimental centipede games conducted by McKelvey and Palfrey (1992). The authors ran experiments using students subjects from Caltech (two sessions) and Pasadena Community College (four sessions). In each subject pool, half the sessions were run on the four-stage game and the other half on the six-stage game. Each experimental session consisted of 18 or 20 subjects, and each subject played the game in the same role either 9 or 10 times. The random matching protocol was such that each player was never matched with another player more than once.

Contrary to backward induction, players did not always take immediately in both four-stage and six-stage games. In fact, a very large majority passed in the first stage. For instance, 94% of the Caltech subjects passed in the first stage in four-stage games (see Tables 3 and 4). The distribution of game outcomes is unimodal with the mode occurring at the intermediate outcomes (outcomes 2 and 3 in four-stage games, outcomes 3 and 4 in six-stage games). These experimental results present a considerable challenge to backward induction.

Tables 3 and 4 also suggest that Caltech subjects take one stage earlier than PCC subjects in both four- and six-stage games. Specifically, the modal outcome is outcome 3 in four-stage games and outcome 4 in six-stage games in the Caltech subject pool, whereas

it is outcome 2 in four-stage games and outcome 3 in six-stage games in the PCC subject pool. These results suggest that the two subject pools exhibit different levels of sophistication.

4.1. Dynamic Level-k Model
We use the dynamic level-k model to explain violations of backward induction in the above data. To facilitate empirical estimation, we restrict attention to a parametric family of distributions for $\phi(\cdot)$. We use the parameter $p_0 \equiv \phi(0)$ to represent the fraction of players with the initial belief that their opponents are nonstrategic thinkers who will play L_0. The remaining $(1 - p_0)$ fraction of players have the initial belief that their opponents are strategic thinkers. We assume that the distribution of this latter type of players $\{\phi(1), \phi(2), \dots, \}$ follows a geometric distribution with parameter q, i.e., $\phi(k) = q \cdot (1 - q)^{k-1}$ for $k = 1, 2, \dots$. The geometric distribution has a natural interpretation: the parameter q is the probability that players believe their opponents will fail to advance to the next higher-level rule. In other words, the higher q is, the less sophisticated players believe their opponents are. The geometric specification is also consistent with the empirical evidence that the modal rule is frequently either L_1 or L_2 (see Costa-Gomes and Crawford 2006; Crawford and Iriberri 2007a, b).

Given the parametric assumptions above, the prediction of the dynamic level-k model for each centipede game and in every game round is completely described by the three model parameters: p_0, q, and β. Note that p_0 and q summarize the heterogeneity in initial beliefs, and β captures each player's initial belief strength. Conditional on the three parameters, the model generates players' choice probabilities for each rule L_k for all game rounds. Let $P_t^i(k \mid p_0, q, \beta)$ denote the model's predicted probability that player i will choose rule L_k in round t.[8] In the first round, $P_1^i(k \mid p_0, q, \beta) = \phi(k - 1)$. For each subsequent round t, there are two possible ways to update the model and generate the predictions:

1. *Open-loop approach*: We assume that subject i mentally simulates and considers all possible sequences of game histories that she could have encountered in all previous rounds 1 to $t - 1$: each sequence $(L^j(1), L^j(2), \dots, L^j(t - 1))$ consists of the rule $L^j(s)$ chosen by the opponent j in round $s = 1, \dots, t - 1$ and occurs with probability $\prod_{s=1}^{t-1} P_s^j(L^j(s) \mid p_0, q, \beta)$. Note that each sequence of observed rules generates a different belief $B^i(t - 1)$

Table 3   **A Comparison of Outcomes in Four-Stage Games Between Caltech and PCC Subjects**

Outcome	4	3	2	1	0
Caltech ($N = 100$)	0.06	0.43	0.28	0.14	0.09
PCC ($N = 181$)	0.08	0.31	0.42	0.16	0.03

Table 4   **A Comparison of Outcomes in Six-Stage Games Between Caltech and PCC Subjects**

Outcome	6	5	4	3	2	1	0
Caltech ($N = 100$)	0.02	0.09	0.39	0.28	0.20	0.01	0.01
PCC ($N = 181$)	0	0.05	0.09	0.44	0.28	0.12	0.02

---

[8] Given the sequential structure of the centipede game, different values of k in L_k may give arise to identical rules. For example, for player B, $L_1 = L_2 = \{-, P, -, T\}$. Such classes of rules can be lumped together in our empirical estimation. For notational convenience, we assume that players always choose the simpler rule (i.e., the lower-level rule).

through the process (2), and thus leads player i to choose a best-response rule in round t according to (3). Therefore, summing over all sequences for which L_k is a best response, we obtain the probability $P_t^i(k \mid p_0, q, \beta)$ of player i choosing rule L_k in round t. This approach allows us to systematically determine each player's choice probabilities in round t using the probabilities that have been calculated for previous rounds $s = 1, \ldots, t-1$.

2. *Closed-loop approach*: Subject i updates her belief using only her own interactions with the matched partners in rounds 1 to $t-1$ (as opposed to simulating and considering all possible such interactions in the population). As a consequence, there is only one sequence (i.e., the observed sample path) consisting of the rule $L^j(s)$ chosen by the opponent j in round $s = 1, \ldots, t-1$, i.e., $(L^j(1), L^j(2), \ldots, L^j(t-1))$. Note that this observed sequence of rules generates a belief $B^i(t-1)$ through the process (2) and thus leads player i to choose a best-response rule in round t according to (3). Given a particular initial belief drawn from distribution $\phi(\cdot)$, our model yields a prediction for the best-response rule in round t for subject i. Therefore, summing over all initial beliefs for which L_k is a best response, we obtain the probability $P_t^i(k \mid p_0, q, \beta)$ of player i choosing rule L_k in round t.

Each approach has its pros and cons. The open-loop approach considers expected frequencies of game plays. It assumes that in each round, individual players not only react to their own histories, but also to all other possible histories that could have occurred if they were to be matched with other players in the population. The closed-loop approach considers individual histories of plays. It assumes that players only respond to their own histories and that future matched players are similar to those encountered in the previous rounds. As a consequence, the open-loop approach may run the risk of not being responsive enough to the observed data, and the closed-loop approach may run the risk of overreacting to the data. Because players face different partners in each round and future matched partners may be quite different from those encountered in the past, players' actual behaviors are likely to fall between these two polar cases. We estimate the model using both approaches to understand which approach is superior in explaining behaviors in the centipede game.

Having specified players' choice probabilities over rules, i.e., $P_t^i(k \mid p_0, q, \beta)$, we now translate these choice probabilities into predicted probabilities over outcomes. Because outcomes are observed in the data, we are interested in the probability of observing a particular outcome in a particular round. Let $P_t(o \mid p_0, q, \beta)$ denote the probability that outcome o will occur in round t, conditional on the model parameters β, p_0, q,

and q. In the centipede game, there is a well-defined prediction over outcomes associated with any pair of rules chosen by the players. For example, if player A chooses L_1 and player B chooses L_3 in a four-stage game, outcome 3 occurs with probability 1; if player A chooses L_0 and player B chooses L_4 in a four-stage game, outcomes 3 and 4 occur with probability 0.5 each. Therefore, based on players' choice probabilities over rules $P_t^i(k \mid p_0, q, \beta)$, we obtain the model's predicted probabilities over outcomes $P_t(o \mid p_0, q, \beta)$.

Similar to backward induction, the dynamic level-k model predicts that some game outcomes (e.g., outcome 0) will occur with zero probability. To facilitate empirical estimation, we need to incorporate an error structure. We use the simplest possible error structure (Crawford and Iriberri 2007a) to avoid specification bias. We assume that there is an error probability $\epsilon > 0$ each of the $(S+1)$ possible outcomes will occur, and with the remaining probability $[1 - \epsilon \cdot (S+1)]$, our model prediction $P_t(o \mid p_0, q, \beta)$ holds. Harless and Camerer (1994) proposed the use of a uniform error rate to model individual choices, and this approach was first used in the level-k specification by Costa-Gomes et al. (2001). Given that we observe outcomes o_t in each round t, the likelihood function is given by

$$L = \prod_t \{[1 - \epsilon \cdot (S+1)] \cdot P_t(o_t \mid p_0, q, \beta) + \epsilon\}. \qquad (4)$$

We fit the dynamic level-k model to the data using maximum likelihood estimation (MLE). We estimate the model using both the open-loop and closed-loop approaches. In addition, we separately estimate the dynamic level-k model for Caltech and PCC subject pools because of their apparent differences in the degree of sophistication (see Tables 3 and 4). However, we use the same set of parameters (p_0, q, β) to fit both the four-stage and six-stage games. The total error probability is also constrained to be the same across both four-stage and six-stage games, i.e., $5 \cdot \epsilon_{(4)} = 7 \cdot \epsilon_{(6)}$, where $\epsilon_{(S)}$ is the error probability used to fit the data of S-stage games. For brevity, we report estimates of $\epsilon = (\epsilon_{(4)} + \epsilon_{(6)})/2$.

Interestingly, the parameter estimates from both the open-loop and closed-loop approaches provide the same qualitative finding: Caltech subjects chose higher-level rules than PCC subjects.[9] The overall likelihood for the open-loop approach, however, was higher than that of the closed-loop approach. The open-loop approach yields log-likelihood scores of

---

[9] The parameter estimates from the closed-loop approach are (1) $\hat{p}_0 = 0.06$ (Caltech) and $\hat{p}_0 = 0.12$ (PCC), (2) $\hat{q} = 0.40$ (Caltech) and $\hat{q} = 0.51$ (PCC), and (3) $\hat{\beta} = 2.58$ (Caltech) and $\hat{\beta} = 2.55$ (PCC). These parameter estimates suggest that Caltech players believe there is a smaller fraction of level-0 opponents and the probability opponents cannot advance to the next rule level is lower in its population.

Table 5 Maximum Likelihood Estimates of Dynamic Level-k Model and Its Special Cases (Caltech Subjects)

Parameter	Backward induction	Naive belief	Static level-k	Full model
p_0	0	1	0.01	0.16
q	0	1	0.52	0.54
β	∞	0.10	∞	0.15
ϵ	0.17	0.16	0.06	0.07
Log-likelihood	-355.7	-353.3	-322.9	-311.7

-311.7 (Caltech) and -514.1 (PCC), and the closed-loop approach yields log-likelihood scores of -350.6 (Caltech) and -560.8 (PCC). Because the open-loop approach fits the data better, we shall focus on its estimation results below.

4.1.1. Estimation Results for Caltech Subjects. Table 5 shows the estimation results for Caltech subjects. As shown, the estimated parameters are $\hat{p}_0 = 0.16$, $\hat{q} = 0.54$, $\hat{\beta} = 0.15$, and $\hat{\epsilon} = 0.07$, and the log-likelihood is -311.7. The estimated $\hat{p}_0$ and $\hat{q}$ suggest that, initially, 16% of the players are level 1, 45% are level 2, 21% are level 3, and the remaining 18% are level 4 or higher. The estimated $\hat{\beta} = 0.15$ suggests a small initial belief weight, which implies that repetition unraveling can begin to occur very quickly.[10]

Tables 6 and 7 compare the model prediction and the actual data for the first and subsequent nine rounds in the four-stage and six-stage games for Caltech subjects. The best fitted dynamic level-k model makes two predictions. First, the model predicts that subjects' behavior in the four-stage games should be "closer" to the backward induction prediction, compared to six-stage games (i.e., limited induction property). Second, the model prescribes that unraveling occurs; in fact, the fitted parameters imply that outcome 1 should not occur after round 1. We suppress the error term (i.e., fix $\epsilon = 0$) to generate the theoretical predictions in Tables 6 and 7. The tables show that the model predictions were roughly consistent with the data in the following ways:

1. In four-stage games, the game ends in the first two stages (i.e., outcomes 3 and 4 occur) approximately 40% of the time. However, in six-stage games, the game ends in the first two stages (i.e., outcomes 5 and 6 occur) only approximately 10% of the time.

2. In four-stage games, the proportions of outcomes 0 and 1 decline after the first round. For example, the proportion of outcome 1 decreases from 20%

---

[10] Nagel and Tang (1998) studied normal-form centipede games and showed that fictitious play models (without initial belief weight β) do not fit their experimental data well. In contrast, our results suggest that dynamic level-k models (which are similar to fictitious play models) with initial belief weight β can capture learning well in extensive-form centipede games. This difference could be due to the inclusion of the β parameter in dynamic level-k models, which effectively allows for a different rate of learning across games.

Table 6 A Comparison of Data and Dynamic Level-k Model Prediction in Four-Stage Games (Caltech Subjects)

| Outcome	4	3	2	1	0
Round 1 data ($N = 10$)	0	0.30	0.30	0.20	0.20
Prediction ($\epsilon = 0$)	0.18	0.25	0.46	0.11	0
Rounds 2–10 data ($N = 90$)	0.07	0.44	0.28	0.13	0.08
Prediction ($\epsilon = 0$)	0.18	0.66	0.15	0	0

Table 7 A Comparison of Data and Dynamic Level-k Model Prediction in Six-Stage Games (Caltech Subjects)

Outcome	6	5	4	3	2	1	0
Round 1 data ($N = 10$)	0	0	0.30	0.10	0.60	0	0
Prediction ($\epsilon = 0$)	0.04	0.06	0.13	0.25	0.42	0.10	0
Rounds 2–10 data ($N = 90$)	0.02	0.10	0.40	0.30	0.16	0.01	0.01
Prediction ($\epsilon = 0$)	0	0.01	0.24	0.62	0.12	0	0

to 13%. In six-stage games, outcomes 0 and 1 rarely occur (approximately 1% of the time).

We also fit three nested cases of the dynamic level-k model. They are all rejected by the likelihood ratio test. The backward induction model (i.e., $p_0 = q = 0, \beta = \infty$) yields a log-likelihood of -355.7, which is strongly rejected in favor of the full model ($\chi^2 = 88.0$). The naive belief model (i.e., $p_0 = q = 1$) (which assumes that all players initially believe that their opponents will always choose randomly) yields an estimated $\hat{\beta} = 0.10$ and a log-likelihood of -353.3. The model is again strongly rejected with $\chi^2 = 83.2$. The static level-k model corresponds to $\beta = \infty$. Because players' initial belief will persist throughout the game plays, they will always choose the same rule L_k across rounds. This restriction yields parameter estimates of $\hat{p}_0 = 0.01$ and $\hat{q} = 0.52$ and provides a reasonable fit with a log-likelihood of -322.3. However, the static model is also rejected by the likelihood ratio test ($\chi^2 = 21.2$, $p < 0.001$).

Given the MLE estimates of $\hat{p}_0 = 0.16$, $\hat{q} = 0.54$, and $\hat{\beta} = 0.15$, we can generate the dynamic level-k model's predicted frequencies for each of the outcomes. Figure 4 shows the actual and predicted frequencies of each outcome. The top panel shows the results for the four-stage game, and the bottom panel shows the results for the six-stage game. Backward induction predicts that only outcome 4 in the four-stage game and only outcome 6 in the six-stage game can occur (i.e., player A takes immediately). This backward induction prediction is strongly rejected by the data. As shown, the dynamic-k model does a reasonable job in capturing the unimodal distribution of the outcomes. In addition, the dynamic level-k model is able to capture the two most frequently occurring outcomes. Specifically, the dynamic level-k model correctly predicts that the two most frequently played outcomes are 2 and 3 in the four-stage game and 3 and 4 in the six-stage game.

Figure 4   Dynamic Level-*k* Model Fit (Dark Bars) and Data (Light Bars) for Caltech Subjects
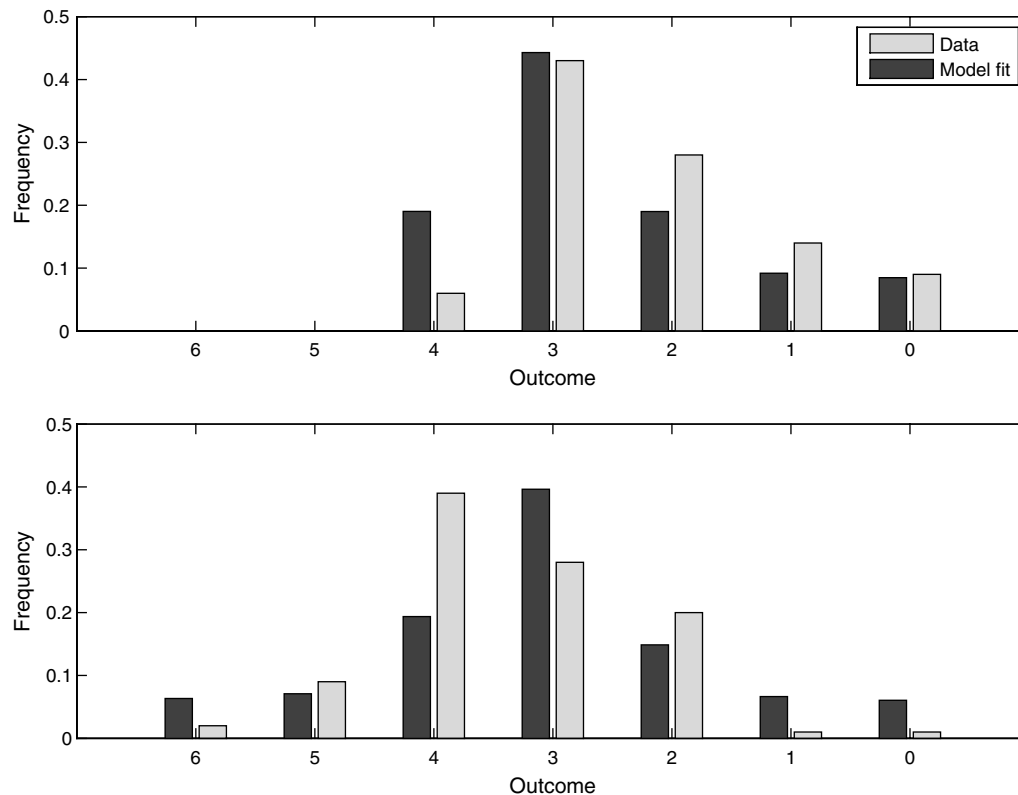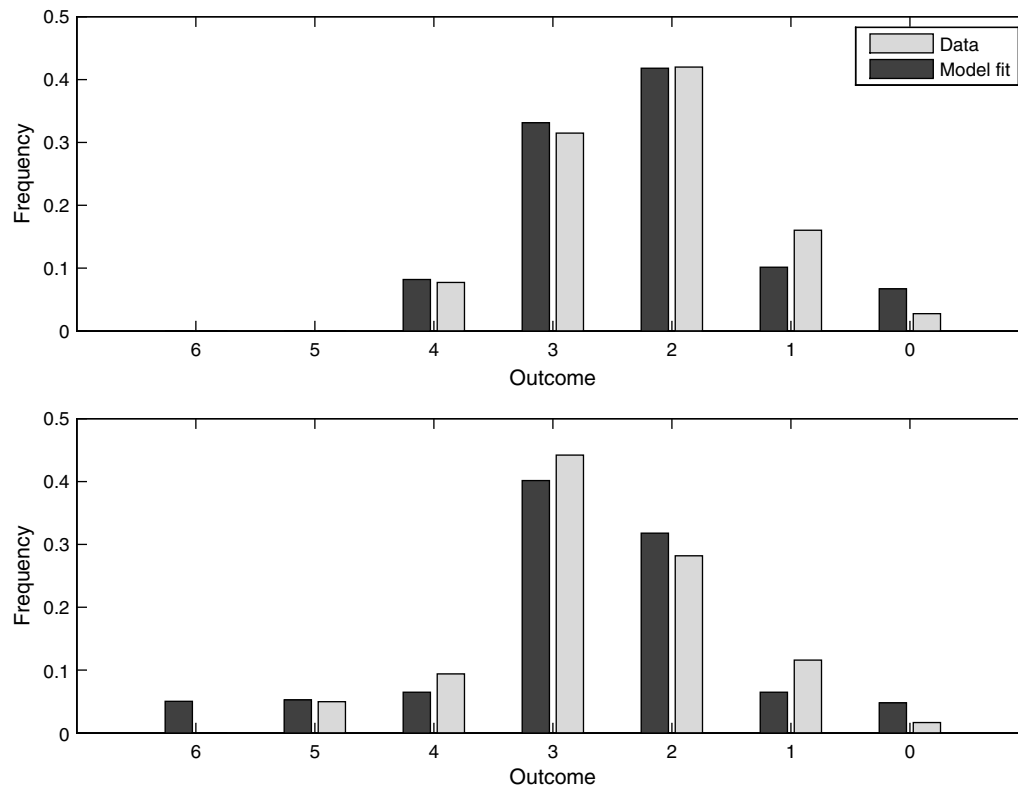

4.1.2.   Estimation Results for PCC Subjects. We also fit our model to the data obtained from PCC subjects. Table 8 shows the estimation results. The dynamic level-*k* model yields parameter estimates of $\hat{p}_0 = 0.26$, $\hat{q} = 0.53$, $\hat{\beta} = 0.80$, and $\hat{\epsilon} = 0.06$. These estimates suggest that, like the Caltech subjects, the PCC subjects also learn from past plays but they are more stubborn (because they have a higher initial belief weight of $\hat{\beta} = 0.80$). The results also show that a larger proportion of PCC subjects play the level-1 rule initially (with $\hat{p}_0 = 0.26$ for PCC subjects, compared to $\hat{p}_0 = 0.16$ for Caltech subjects). The naive belief model ($p_0 = q = 1$) gives a log-likelihood score of -599.7, and the backward induction prediction ($p_0 = q = 1$, $\beta = \infty$) gives a log-likelihood score of -643.8, so both models are strongly rejected by the likelihood ratio test. Like before, the static level-*k* model is strongly rejected by the likelihood ratio test ($\chi^2 = 22.4$, $p < 0.01$).

Given the MLE estimates of $\hat{p}_0 = 0.26$, $\hat{q} = 0.54$, and $\hat{\beta} = 0.80$ for PCC subjects, we can generate the dynamic level-*k* model's predicted frequencies for each of the outcomes. Figure 5 shows the actual and predicted frequencies of each outcome. The top panel shows the results for the four-stage game, and the bottom panel shows the results for the six-stage game. The plot shows that the dynamic level-*k* model provides a reasonably good fit of the data for PCC subjects.

Finally, we show that our best fitted dynamic level-*k* model is indeed consistent with the limited induction and repetition unraveling properties described in §2. Following the approach described in §2, we calculate the deviation from backward induction for each outcome using Equation (1) and then take expectations using the predicted frequencies of the best fitted model. For Caltech subjects, the expected deviation from backward induction is 0.36 for the four-stage game and increases to 0.49 for the six-stage game. Similarly, for PCC subjects, the expected deviation from backward induction is lower in the four-stage game than in the six-stage game (0.44 versus 0.55). These results support the limited induction property. Next, using the best fitted models for the two subject pools, we find that the expected deviation from backward induction for Caltech subjects is higher in the first five rounds than in the last

Table 8   Maximum Likelihood Estimates of Dynamic Level-*k* Model and Its Special Cases (PCC Subjects)

Parameter	Backward induction	Naive belief	Static level *k*	Full model
p_0	0	1	0.19	0.26
q	0	1	0.59	0.53
β	∞	0.10	∞	0.80
ϵ	0.17	0.13	0.02	0.06
Log-likelihood	-643.8	-599.7	-525.3	-514.1

Figure 5 **Dynamic Level-*k* Model Fit (Dark Bars) and Data (Light Bars) for PCC Subjects**

five rounds (0.43 versus 0.41; these figures are averaged over four-stage and six-stage games with equal probability); the same inequality holds separately for both four-stage and six-stage games. Similarly, for PCC subjects, the expected deviation also decreases from 0.51 in the first five rounds to 0.47 in the last five rounds. These results suggest that the repetition unraveling property holds. Interestingly, note that PCC subjects exhibit a higher deviation than the Caltech subjects. In summary, we conclude that the dynamic level-*k* model can account for limited induction and repetition unraveling properties, the two systematic violations of backward induction.

4.2. Reputation-Based Model

As indicated before, backward induction does not admit passing behavior in centipede games. To explain why players pass, Kreps et al. (1982) proposed transforming a centipede game into a game of incomplete information by introducing some uncertainty over the player type. If there is a small fraction θ of players that always pass (i.e., the altruists), then even the self-interested players may find it in their best interest to pass initially. However, this reputation-based model generates the following predictions that run counter to the observed data (McKelvey and Palfrey 1992):

1. When θ is small ($0 \leq \theta \leq 1/49$), the model will not correctly predict the most frequently occurring

outcomes. For instance, the model predicts the modal outcome to be outcome 4 in four-stage game, whereas the data suggest it should be outcome 2. Similarly, the model predicts the two most frequently occurring outcomes to be outcomes 4 and 5 in six-stage game, whereas the data suggest they should be outcomes 2 and 3.

2. When θ is intermediate ($1/49 < \theta \leq 1/7$), the model will not admit outcome 4 in the four-stage game and outcomes 4–6 in the six-stage game. Yet, the occurrences of these outcomes are not negligible in the data, particularly for Caltech subjects.

3. When θ is large ($\theta > 1/7$), the model predicts players either pass in every stage (outcome 0) or take at the last stage (outcome 1), which is contrary to the data.

The above discussion suggests that there exists no common θ value that will explain the unimodal distribution of game outcomes in both the four-stage and six-stage games well.

We estimate the reputation-based model using the same data. Like before, because the reputation-based model predicts that some outcomes will occur with zero probability, we add an error probability ϵ to every possible outcome. The best fitted model yields $\hat{\theta} = 0.05$ and $\hat{\epsilon} = 0.11$ for Caltech subjects and $\hat{\theta} = 0.08$ and $\hat{\epsilon} = 0.05$ for PCC subjects. The log-likelihood scores are -329.8 and -518.8, respectively. Finally,

we compare the reputation-based model with the dynamic level-k model using the standard Akaike information criterion (AIC) because the former is not nested by the latter. The AIC of the dynamic level-k model is 631.4 and 1,036.2 for the Caltech and PCC subjects, respectively. Both figures are lower than the corresponding figures of 663.6 and 1,041.6 for the reputation-based model. These results suggest that the dynamic level-k model does a better job of fitting the data compared to the reputation-based model. It is possible to consider a dynamic generalization of the reputation-based model, in which the self-interested players learn about the fraction of altruists (i.e., the θ parameter) over time and update their mixed strategy accordingly. Such a model is likely to provide a better fit. However, given an observed outcome, it is not clear how players should revise their beliefs of θ. In contrast, in the dynamic level-k model, because players observe opponents' rules, the belief updating is more straightforward.

4.3. Models of Social Preferences

Could the unimodal distribution of game outcomes in centipede games be attributed to social preferences? To address this question, we apply the inequity aversion model of Fehr and Schmidt (1999). Let x_i and x_{-i} denote the material payoffs of player i and opponent $-i$, respectively. Then, the payoff of player i, $U_i(x_i, x_{-i})$, is given by

$$U_i(x_i, x_{-i}) = x_i - \alpha \cdot [x_i - x_{-i}]^+ - \beta \cdot [x_{-i} - x_i]^+, \quad (5)$$

where $0 < \alpha < 1$ captures a player's aversion to being ahead, and $\beta > 0$ captures a player's aversion to being behind. The parameter α is between 0 and 1 because a player who is ahead will not give up a \$1 to benefit her opponent by less than \$1. One can then solve the game by backward induction given these revised payoff functions.

THEOREM 3. *In the centipede game, social preferences lead to either taking immediately or passing all the way in both the four-stage and six-stage games. Specifically, social preferences predicts that if $3 \cdot \alpha - 6 \cdot \beta < 2$, players will take immediately; otherwise they will pass all the way.*

PROOF. See Appendix A.

The above theorem states that a model of inequity aversion admits only either outcome 0 or 4 in the four-stage game and either outcome 0 or 6 in the six-stage game. The occurrence of either outcome depends on the fairness parameters α and β. If players are sufficiently averse to being behind (high β), they will take immediately, which is consistent to backward induction. On the other hand, if players

are sufficiently averse to being ahead (high α), they will pass all the way. Both predictions are inconsistent with the unimodal distribution of actual game outcomes.

We estimate the social preference model using the centipede game data. If $3 \cdot \alpha - 6 \cdot \beta < 2$, the social preference model yields identical predictions as backward induction, so the best fit yields a log-likelihood of -355.7 for Caltech subjects and -643.8 for PCC subjects. If $3 \cdot \alpha - 6 \cdot \beta \geq 2$, the social preference model predicts that players pass all the way, and we find that best fit again yields a log-likelihood of -355.7 and -643.8. Hence the model performs poorly compared to the dynamic level-k model. However, the utility function in (5) can be generalized to capture nonlinear social preferences. Such extensions may describe the data better by admitting intermediate game outcomes and can be explored in future research.

5. Sequential Bargaining Games

This section demonstrates the generalizability of the dynamic level-k model to accommodate other violations of backward induction. For example, such violations have been studied in sequential bargaining games (Guth et al. 1982), the chain-store paradox (Selten 1978), and the finitely repeated prisoners' dilemma (Kreps et al. 1982). In particular, we apply the dynamic level-k model to study sequential bargaining games.

This section investigates the generalizability question by applying the same dynamic level-k model to sequential bargaining games (Stahl 1972, Rubinstein 1982). Like centipede games, sequential bargaining games consist of a finite number of stages. In each stage, two players bargain over a fixed pie, with one player proposing a division of the pie and the other player either accepting or rejecting the proposed offer. Upon acceptance, the game ends, and both players receive their shares according to the proposal. Upon rejection, the game proceeds to the next stage, and it is now the other player's turn to propose a division. However, in each subsequent stage, the size of the pie shrinks. In the final stage, both players receive nothing if the final proposal is rejected. The backward induction principle provides a sharp prediction for a sequential bargaining game: in the very first stage, the proposer will demand the largest possible share that will be accepted by the responder and the game ends. Camerer et al. (1993) provided some of the earliest experimental evidence to show that subjects rarely look at crucial payoff information that is needed to apply backward induction in sequential bargaining games. Johnson et al. (2002) built on this work by showing that these violations of backward induction cannot be explained by social preferences

Table 9   Initial Demands by Player A for Rounds 1–10 and Rounds 11–20 in Binmore et al. (2002)

	Disagreement payoffs		
	(10, 10)	(70, 10)	(10, 60)
Rounds 1–10	64.9 (10.3)	82.8 (8.2)	39.8 (9.8)
Rounds 11–20	66.8 (6.6)	83.4 (3.9)	36.0 (6.9)

Table 10   Initial Demands by Level-k Player A with Different Disagreement Payoffs

	Disagreement payoffs		
k	(10, 10)	(70, 10)	(10, 60)
0	Random	Random	Random
1	55	85	55
≥ 2	90	90	40

because subjects make similar offers to robot players. Overall, both studies provide compelling "mouselab" process data to show that subjects did not apply backward induction in these games.

Binmore et al. (2002) investigated the simplest possible sequential bargaining games called the ultimatum games (refer to seminal work by Guth et al. (1982) and Guth and Tietz (1990) for a comprehensive review of the literature). In ultimatum games, the first player (player A) proposes a division of a pie of size $\pi = 100$, and the game ends with the second player (player B) either accepting or rejecting the proposed offer. In their experimental design, the authors consider different disagreement payoffs, i.e., payoffs players A and B would receive (V_A, V_B), respectively, if player A's proposal were rejected. Three disagreement payoffs were tested, with (V_A, V_B) being (10, 10), (70, 10), or (10, 60). Furthermore, subjects played the game repeatedly for 20 rounds. The repeated plays allow us to detect any shift in initial demands by player A over time in these ultimatum games.

Table 9 presents player A's demands in these experiments. We group and compare data for the first 10 rounds and data for the last 10 rounds to detect learning over time. We note that player A's initial demands increase over game rounds when disagreement payoff vectors of (10, 10) and (70, 10) are used. For the former, the increase from 64.9 to 66.8 is highly significant ($t = -4.50$, p-value < 0.001). For the latter, the increase from 82.8 to 83.4 is marginally significant at 10% level ($t = -1.90$, p-value $= 0.057$). On the other hand, player A's initial demands decrease from 39.8 to 36.0 over game rounds when disagreement payoffs of (10, 60) are used, and this difference is highly significant ($t = 9.02$, p-value < 0.001).

Can the dynamic level-k model predict these shifts in initial demands over time, especially when the direction of shift varies across the different disagreement payoff vectors (V_A, V_B)? The answer is yes. Let us explain why. We first need to construct the dynamic level k's rule hierarchy. Like before, level-0 players are assumed to choose a number between 0 and 100 randomly. For player A, this randomly chosen number is the initial demand, and for player B, this number is the acceptance threshold (i.e., only player A's offer that is above this number will be accepted). Next, level-1 players best respond to

level-0 players' randomization strategies. A level-1 player B will accept any offer greater than V_B. A level-1 player A's offer of y will be accepted (by a level-0 player B) with probability $y/100$, and thus yields an expected payoff of $(y/100)(100 - y) + ((100 - y)/100)V_A$. Hence, the optimal offer is $y^* = (100 - V_A)/2$, so the optimal demand is $(100 + V_A)/2$. Finally, we consider level-2 players' optimal best responses. As above, a level-2 player B will accept any offer greater than V_B. A level-2 player A will offer the minimum acceptable amount to a level-1 player B, which is V_B. Thus, the initial demand by a level-2 player A is $100 - V_B$. Plugging in the values for (V_A, V_B) used in the experiments, we obtain each rule level's optimal best response in Table 10 (note that L_k for $k > 2$ has the same optimal best response as L_2).

As discussed above, the dynamic level-k model suggests that all players will eventually advance to higher-level rules and play the backward induction outcome. Specifically, we can show that lower-level rules will be successively eliminated as players gather more observations of their opponents' rules through repeated game plays. For ultimatum games with disagreement payoffs, L_1 players will eventually play the strategies for L_2 and above. This transition leads to an increase in initial demands when the disagreement payoffs are (10, 10) or (70, 10) and a decrease in initial demands when the disagreement payoffs are (10, 60). This prediction indeed matches the direction of the shift in initial demands observed in the experiments of Binmore et al. (2002).

In a related study, Neelin et al. (1988) investigated a two-stage sequential bargaining game with pie sizes of $5.00 (Stage I) and $1.25 (Stage II) and a three-stage sequential bargaining game with pie sizes of $5.00 (Stage I), $2.50 (Stage II), and $1.25 (Stage III). Note that the pie shrinks faster in the two-stage than in the three-stage game. The pie sizes are chosen so that the backward induction prediction is identical across the two-stage and three-stage games; i.e., in the very first stage, the proposer will propose to keep $3.75, and the responder will accept and receive $1.25. However, experimental results reveal a stark disparity in initial demands between the two-stage and three-stage games. In the two-stage game, the modal amount that is demanded by the proposer in the first

Table 11 Initial Demand Chosen by Level-*k* Player in Two-Stage and Three-Stage Sequential Bargaining Games

k	Two-stage games	Three-stage games
0	Random	Random
1	3.75	2.66
2	4.69	3.75
3	3.75	2.81
≥ 4	3.75	3.75

round is \$3.75 (chosen by 15 out of 40 subjects). However, in the three-stage game, the modal demand is \$2.50 (chosen by 28 out of 40 subjects). The mean initial demands for the two games are statistically different, even though backward induction predicts no difference between them. We investigate whether the same dynamic level-*k* model can explain the stark difference in initial demands between the two-stage and three-stage sequential bargaining games.

Again, we construct a level-*k* rule hierarchy to determine the initial demand chosen by a level-*k* player in the first stage as follows. Like before, we assume that level-0 players choose randomly within the feasible strategy space at each stage. For each higher-level rule $L_k, k \geq 1$, we can iteratively calculate its best response like before (see Appendix B for details). Table 11 summarizes each rule level's initial demand prediction.

Notice that rules L_1, L_2, and L_3 all prescribe a *smaller* initial demand in the three-stage game compared to the two-stage game. Also, recall that under the dynamic level-*k* model, all players choose rules L_1 and above. Therefore, if the proportion of players choosing rules L_1, L_2, and L_3 is large enough (i.e., $\phi(0), \phi(1)$, and $\phi(2)$ are high enough), then the dynamic level-*k* model will indeed predict a lower initial demand in three-stage games, as observed in the Neelin et al. (1988) sequential bargaining experiments.

6. Conclusions

In experimental centipede games, backward induction is frequently violated. We develop a dynamic level-*k* model to explain two systematic violations of backward induction. First, players tend to deviate more from backward induction in games with a greater number of stages or subgames. Our model captures this limited induction phenomenon by allowing players to have heterogeneous initial beliefs about others' level of sophistication and hence adopt different rules from an iteratively defined rule hierarchy. As a consequence, if a player adopts the same rule, there is always a higher deviation from backward induction in a game with more stages. Second, players move closer to backward induction with

repetition. Our model captures the repetition unraveling phenomenon by allowing players to update beliefs of their opponents' rules and hence adjust their own rules over repetition. Unraveling occurs because the dynamic level-*k* model implies a domino effect over repetition as lower-level rules are successively eliminated.

We fit our model to the experimental centipede games of McKelvey and Palfrey (1992). Our estimation results show that the dynamic level-*k* model captures the unimodal distribution of game outcomes in both four-stage and six-stage games reasonably well. Special cases including backward induction and the static level-*k* model are strongly rejected by the data. Interestingly, Caltech subjects, who are arguably more sophisticated, believe their opponents are more sophisticated, and learn faster than PCC subjects.

We apply the same dynamic level-*k* model to account for violations of backward induction in sequential bargaining games. Our results suggest that the dynamic level-*k* model is capable of capturing systematic violations of backward induction in two separate studies and looks promising as a possible empirical alternative in predicting behaviors in these games.

Our model is a generalization of backward induction. In addition, the dynamic level-*k* model converges to backward induction in the limit, so it can be conceptualized as a tracing procedure for backward induction, and as a consequence, the former provides a dynamic foundation for the latter. Framing it this way, the violations of backward induction in the experiments are simply "transient" behaviors that the dynamic level-*k* model explicitly characterizes.

Appendix A. Proofs

PROOF OF THEOREM 1. Consider two centipede games with s and S subgames, where $s < S$. Suppose that the proportion of players who choose each rule level is the same in both games. Let π_k denote the proportion of players who choose rule L_k, where $k \geq 1$. Let $E\delta(S)$ denote the expected deviation in a game with S subgames.

Suppose player A adopts rule L_{k_A} and player B adopts rule L_{k_B} in a game with S stages. Then, the deviation is

$$\delta(L_{k_A}, L_{k_B}, S) = \frac{1}{S}[S - (\min(\lfloor k_A/2 \rfloor, S/2) + \min(\lceil k_B/2 \rceil, S/2))],$$

which is weakly increasing in S for every k_A and k_B.

Now, comparing the expected deviations in both games, we have

$$E\delta(S) - E\delta(s)$$
$$= \sum_{k_A=1}^{S} \sum_{k_B=1}^{S} \pi_{k_A} \pi_{k_B} [\delta(L_{k_A}, L_{k_B}, S) - \delta(L_{k_A}, L_{k_B}, s)] \geq 0,$$

which gives the required result. $\square$

PROOF OF THEOREM 2. Consider any player in the centipede game. For concreteness, we focus on the four-stage game (see Figure 1) but the proof proceeds in the same way for any number of stages. Suppose player i holds beliefs $B^i(t) = (B_0^i(t), B_1^i(t), \ldots, B_4^i(t))$ at time t, where $B_k^i(t)$ denotes the player's belief of the probability that the opponent will use rule L_k and is calculated from rule counts $N^i(t) = (N_0^i(t), N_1^i(t), \ldots, N_4^i(t))$ as in (2). Given the above beliefs, the player chooses a best-response rule according to (3).

Notice that player A may choose from the rules L_1, L_3 (which is equivalent to L_2), and L_4, and player B may choose from the rules L_2 and L_4 (which are equivalent to L_1 and L_3, respectively). Using the payoffs in Figure 1, we calculate the expected payoffs of each player from using each rule. For player A, the expected payoffs from choosing L_1, L_3, and L_4 are, respectively, $19b_0^A + 8b_{1,2}^A + 2b_{3,4}^A$, $9b_0^A + 16b_{1,2}^A + 2b_{3,4}^A$, and $4b_0^A + 4b_{1,2}^A + 4b_{3,4}^A$, where b_R^i denotes player i's beliefs of the opponent choosing a rule level from R. For example, at time t, $b_{1,2}^A = b_1^A(t) + b_2^A(t)$. Similarly, for player B, the expected payoffs from choosing rules L_1 and L_3 are, respectively, $9.5b_0^B + 32b_1^B + 4b_{2,3}^B + b_4^B$ and $4.5b_0^B + 8b_1^B + 8b_{2,3}^B + b_4^B$.

In the model, each player starts at $t = 0$ with a belief weight of β on some k and adds a unit weight in each round t. Note that no player will choose L_0, so the weights will be added to levels $k \geq 1$ in each round.

Consider player A. For L_1 to be a best response, it must dominate L_2 and L_4, so we must have $10b_0^A \geq 8b_{1,2}^A$ and $15b_0^A + 4b_{1,2}^A \geq 2b_{3,4}^A$, which implies $10b_0^A \geq b_{3,4}^A$. Because weights will only be added to levels $k \geq 1$, at least one of the above inequalities will not hold after $R_1 \equiv \lfloor 10\beta + (5/4)\beta \rfloor + 1 = \lfloor 11.25\beta \rfloor + 1$ rounds. Thereafter, player A will never choose L_1.

Now, consider player B. For L_2 (or L_1) to be a best response, it must dominate L_4 (or L_3), implying that $5b_0^B + 16b_1^B \geq 4b_{2,3}^B$. After R^A rounds (calculated above), player A never chooses L_1, so all weights will be added to levels $k \geq 2$ subsequently. Now, after another $R_2 \equiv \lfloor (5\beta + 16R_1)/4 \rfloor + 1$ rounds, the above inequality will not hold and player B will never choose L_2.

Finally, we move back to player A. For L_3 (or L_2) to be a best response, it must dominate L_4, requiring $5b_0^A + 8b_{1,2}^A \geq 2b_{3,4}^A$. After $R_1 + R_2$ rounds (calculated above), player B never chooses L_2, so all weights will be added to levels $k \geq 3$. After an additional $R_3 \equiv \lfloor (5\beta + 8(R_1 + R_2))/2 \rfloor + 1$ rounds, the inequality above will not hold, and player A will never choose L_3. Therefore, after $R_1 + R_2 + R_3$ rounds,

both players will always choose $L_4 = L_\infty$, and the backward induction outcome always occurs. $\square$

PROOF OF THEOREM 3. We will prove the result for a four-stage game. The proof for a six-stage game proceeds similarly. First, consider the case where $3\alpha - 6\beta \leq 2$. In the last stage, the payoffs to players A and B, respectively, are $(8, 32)$ if player B takes and $(64, 16)$ if player B passes. The corresponding utilities to player B are $32 - 24\alpha$ and $16 - 48\beta$, respectively. Because the condition $3\alpha - 6\beta \leq 2$ holds, player B prefers to take. Reasoning backward, in the second last stage, the payoffs to both players are $(16, 4)$ if player A takes and $(8, 32)$ if player A passes (because player B will then take in the last stage). The corresponding utilities to player A are $16 - 12\alpha$ and $8 - 24\beta$, so by the same condition above, player A takes. Using the same reasoning, players will take in every stage.

Next, consider the other case where $3\alpha - 6\beta \geq 2$. By the calculations above, we know that player B will now pass in the last stage. Now, in the second last stage, the payoffs to players A and B are $(16, 4)$ if player A takes and $(64, 16)$ if player A passes (because player B will also pass in the last stage). The utilities are $16 - 12\alpha$ and $64 - 48\alpha$. Because $\alpha \leq 1$, the latter is greater, and thus player A passes. Proceeding backward, in the third last stage (i.e., Stage II), player B's payoff from taking is $8 - 6\alpha$, which is less than his payoff $32 - 24\alpha$ from taking in the last stage, because $\alpha \leq 1$. Because player B prefers passing in the last stage, the same applies in this stage as all players will pass subsequently. Using the same logic, players will pass in every stage. $\square$

Appendix B. Calculations for Bargaining Games

We first consider the two-stage and three-stage sequential bargaining games studied by Neelin et al. (1988). Player A first acts as the proposer follower by player B. We assume that level-0 players randomize uniformly. Specifically, in the first stage, the level-0 player A randomly chooses a number between 0 and 5.00 as the initial demand. Similarly, in the second stage, the level-0 player B randomly chooses a number between 0 and 2.50, which serves both as an acceptance threshold and a demand: if player A's offer from Stage I exceeds this number, player B accepts the offer; otherwise, player B rejects and demands this amount. Next, in the third stage, the level-0 player A randomly chooses a number between 0 and 1.25 as the acceptance threshold and demand. Finally, when responding to this final offer from the third stage, the level-0 player B randomly chooses a number between 0 and 1.25 as the acceptance threshold. (For convenience, we shall refer to this final decision as Stage III+.) Given this specification for L_0 strategies, we can iteratively calculate the strategies for L_1 and above.

Now, we consider the L_1 player. In Stage III+, player B best responds and thus will accept any positive offer. In Stage III, upon rejecting the previous offer, player A must choose some amount $y \in [0, 1.25]$ to offer, which is accepted with probability $y/1.25$ and thus yields payoff $(y/1.25) \cdot (1.25 - y)$; the optimal offer is thus 0.625, and the expected payoff is 0.3125, which is thus the acceptance threshold. Now, we move to Stage II. Here, upon rejecting the previous offer, player B must choose some amount $y \in [0, 2.5]$ to offer; any offer of 1.25 or above will be accepted with

certainty and is dominated by offering 1.25, so we focus on $y \in [0, 1.25]$. The offer y will be accepted with probability $y/1.25$ and rejected with probability $(1.25 - y)/1.25$, and upon rejection, B will accept A's subsequent (final) offer. Thus, B's expected payoff from offering y to A is given by $(y/1.25)(2.5 - y) + (1.25 - y)/1.25(1.25 - (1.25 + y)/2)$, which attains a maximum value of 1.25 at $y^* = 1.25$. So, B's acceptance threshold at this stage is 1.25. Finally, we move to Stage I. Here, player A makes an offer $y \in [0, 5]$ to player B. Similar to above, we focus on $y \in [0, 2.5]$ because offers above 2.5 are dominated by an offer of 2.5. Following a similar logic as above, player A's payoff function is

$$
\pi(y) = \begin{cases} \dfrac{y}{2.5}(5-y) + \dfrac{2.1875 - y}{2.5}\left(2.5 - \dfrac{2.1875 + y}{2}\right) \\ \quad + \dfrac{2.5 - 2.1875}{2.5}(0.3125) \qquad y \le 2.1875, \\ \dfrac{y}{2.5}(5-y) + \dfrac{2.5 - y}{2.5}(0.3125) \quad y > 2.1875, \end{cases}
$$

which is maximized at $y^* = 2.34375$. Thus, the level-1 player A makes an initial demand of $5 - y^* = 2.66$. Next, we move to the L_2 player. In Stage III+, player B will accept any positive offer. In Stage III, player A will demand 1.25 because the level-1 player A will accept any positive offer. In Stage II, player B will demand $2.5 - 0.3125 = 2.1875$ because the level-1 player A's acceptance threshold (in Stage III) is 0.3125. Finally, in Stage I, player A will demand $5 - 1.25 = 3.75$ because the level-1 player B's acceptance threshold (in Stage II) is 1.25. The same reasoning gives the strategies for higher-level players.

References

Aumann R (1995) Backward induction and common knowledge of rationality. *Games Econom. Behav.* 8(1):6–19.

Binmore K, McCarthy J, Ponti G, Samuelson L, Shaked A (2002) A backward induction experiment. *J. Econom. Theory* 104:48–88.

Brown GW (1951) Iterative solutions of games by fictitious play. Koopmans TC, ed. *Activity Analysis of Production and Allocation* (Wiley, New York), 374–376.

Camerer C, Ho T-H (1999) Experience-weighted attraction learning in normal form games. *Econometrica* 67(4):837–874.

Camerer CF, Ho T-H, Chong J-K (2004) A cognitive hierarchy model of games. *Quart. J. Econom.* 119(3):861–898.

Camerer C, Johnson EJ, Sen S, Rymon T (1993) Cognition and framing in sequential bargaining for gains and losses. Binmore K, Kirman A, and Tani P, eds. *Frontiers of Game Theory* (MIT Press, Cambridge, MA), 27–48.

Charness G, Levin D (2009) The origin of the winner's curse: A laboratory study. *Amer. Econom. J.: Microeconomics* 1(1):207–236.

Costa-Gomes M, Crawford V (2006) Cognition and behavior in two-person guessing games: An experimental study. *Amer. Econom. Rev.* 96(5):1737–1768.

Costa-Gomes M, Weizsacker G (2008) Stated beliefs and play in normal-form games. *Rev. Econom. Stud.* 75(3):729–762.

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

Crawford V, Iriberri N (2007a) Fatal attraction: Salience, naivete, and sophistication in experimental hide-and-seek games. *Amer. Econom. Rev.* 97(5):1731–1750.

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

Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. *Quart. J. Econom.* 114(3):817–868.

Fey M, McKelvey RD, Palfrey TR (1996) An experimental study of constant-sum centipede games. *Internat. J. Game Theory* 25(3):269–287.

Fudenberg D, Levine D (1998) *Theory of Learning in Games* (MIT Press, Cambridge, MA).

Guth W, Tietz R (1990) Ultimatum bargaining behavior a survey and comparison of experimental results. *J. Econom. Psych.* 11(3):417–449.

Guth W, Schmittberger R, Schwarze B (1982) An experimental analysis of ultimatum bargaining. *J. Econom. Behav. Organ.* 3(4):367–388.

Harless DW, Camerer C (1994) The predictive utility of generalized expected utility theories. *Econometrica* 62(6):1251–1289.

Ho T-H, Camerer C, Chong J-K (2007) Self-tuning experience-weighted attraction learning in games. *J. Econom. Theory* 133(1):177–198.

Ho T-H, Camerer C, Weigelt K (1998) Iterated dominance and iterated best response in experimental *P*-beauty contests. *Amer. Econom. Rev.* 88(4):947–969.

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

Kreps DM, Milgrom P, Roberts J, Wilson R (1982) Rational cooperation in the finitely repeated prisoners' dilemma. *J. Econom. Theory* 27(2):245–252.

Levitt SD, List JA, Sadoff SE (2011) Checkmate: Exploring backward induction among chess players. *Amer. Econom. Rev.* 101(2): 975–990.

McKelvey RD, Palfrey TR (1992) An experimental study of the centipede game. *Econometrica* 60(4):803–836.

Nagel R (1995) Unraveling in guessing games: A experimental study. *Amer. Econom. Rev.* 85(5):1313–1326.

Nagel R, Tang F (1998) Experimental results on the centipede game in normal form: An investigation in learning. *J. Math. Psych.* 42(2–3):356–384.

Neelin J, Sonnenschein H, Spiegel M (1988) A further test of non-cooperative bargaining theory: Comment. *Amer. Econom. Rev.* 78(4):824–836.

Palacios-Huerta I, Volij O (2009) Field centipedes. *Amer. Econom. Rev.* 99(4):1619–1635.

Rosenthal RW (1981) Games of perfect information, predatory pricing and the chain-store paradox. *J. Econom. Theory* 25(1):92–100.

Rubinstein A (1982) Perfect equilibrium in a bargaining model. *Econometrica* 50(1):97–109.

Selten R (1978) The chain store paradox. *Theory Decision* 9(2): 127–159.

Stahl D (1996) Boundedly rational rule learning in a guessing game. *Games Econom. Behav.* 16(2):303–330.

Stahl D, Wilson P (1995) On players' models of other players—Theory and experimental evidence. *Games Econom. Behav.* 10(1): 218–254.

Stahl I (1972) *Bargaining Theory* (Stockholm School of Economics, Stockholm).