

NBER WORKING PAPER SERIES

CHECKMATE: EXPLORING BACKWARD INDUCTION AMONG CHESS PLAYERS

Steven D. Levitt  
John A. List  
Sally E. Sadoff

Working Paper 15610  
<http://www.nber.org/papers/w15610>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
December 2009

We would like to thank Martin Dufwenberg, Philip Reny, five anonymous referees, and especially, the editor Vince Crawford for insightful comments that improved the study. Trevor Gallen and Elizabeth Sadoff provided truly outstanding research assistance on the ground. Min Sok Lee, Lint Barrage, Tova Levin, Nicholas Simmons and Yana Peysakhovich also provided able research assistance. Affiliations: Steven Levitt and John List: University of Chicago, Department of Economics and National Bureau of Economic Research; Sally Sadoff: University of Chicago, Department of Economics. Mailing address: 1126 E. 59th Street, Chicago, IL 60637. Please direct all correspondence to Sally Sadoff. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2009 by Steven D. Levitt, John A. List, and Sally E. Sadoff. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Checkmate: Exploring Backward Induction Among Chess Players  
Steven D. Levitt, John A. List, and Sally E. Sadoff  
NBER Working Paper No. 15610  
December 2009  
JEL No. C9,C91,C92,C93,D01

### **ABSTRACT**

Although backward induction is a cornerstone of game theory, most laboratory experiments have found that agents are not able to successfully backward induct. Much of this evidence, however, is generated using the Centipede game, which is ill-suited for testing the theory. In this study, we analyze the play of world class chess players both in the centipede game and in another class of games – Race to 100 games – that are pure tests of backward induction. We find that world class chess players behave like student subjects in the centipede game, virtually never playing the backward induction equilibrium. In the race to 100 games, in contrast, we find that many chess players properly backward induct. Consistent with our claim that the Centipede game is not a useful test of backward induction, we find no systematic within-subject relationship between choices in the centipede game and performance in pure backward induction games.

Steven D. Levitt  
Department of Economics  
University of Chicago  
1126 East 59th Street  
Chicago, IL 60637  
and NBER  
slevitt@midway.uchicago.edu

Sally E. Sadoff  
Department of Economics  
University of Chicago  
sadoff@uchicago.edu

John A. List  
Department of Economics  
University of Chicago  
1126 East 59th  
Chicago, IL 60637  
and NBER  
jlist@uchicago.edu

## **I. Introduction**

It is difficult to overstate the profound impact that game theory has had on the economic approach and on the sciences more generally. For that reason, understanding how closely the assumptions that underpin game theoretic analysis conform to actual human decision-making is a question of first-order importance to economists. In this spirit, backward induction represents one of the most basic concepts in game theory. Backward induction played a prominent role in Reinhard Selten's (1965) development of perfect equilibrium, and it has helped to shape the modern refinement literature. Although backward induction is a cornerstone of game theory, existing empirical evidence suggests that economic agents engage in backward induction less frequently than theorists might hope.

Backward induction has fared especially poorly in the centipede game, which was introduced by Robert W. Rosenthal (1981) and has since been extensively analyzed (Kenneth G. Binmore, 1987; Philip J. Reny, 1988; Robert J. Aumann, 1988; David M. Kreps, 1990; Geir B. Asheim and Martin Dufwenberg, 2003). The original centipede game is a two-player, finite-move game in which the subjects alternate choosing whether to end the game or to pass to the other player. The subject's payoff to ending the game at a particular node is greater than the payoff he receives if the other player ends the game at the next node, but less than the payoff earned if the other player elects not to end the game. The player making the final choice gets paid more from stopping than from passing, and thus would be expected to stop. If the opponent will stop at the last node, then conditional on reaching the penultimate node, the player maximizes his earnings by stopping at that node. Following this logic further, backward induction leads to the unique subgame perfect equilibrium: the game is stopped at the first node.

As pointed out in prior research (Rosenthal, 1981, Aumann, 1992, Richard D. McKelvey and Thomas R. Palfrey, 1992, Mark Fey, McKelvey, and Palfrey, 1996, and Klaus G. Zauner, 1999), there are many reasons why players might take actions in the centipede game that diverge from that prescribed by backward induction. Players may face an aversion to the loss of a potential surplus. They may have social preferences for fairness, altruism, or cooperation; or, they may believe that enough other players in the population have these preferences that continuing the game becomes the optimal rational strategy (Aumann, 1995). Similarly, there may be enough players in the population who make backward induction errors that continuing the game becomes the optimal rational strategy. Reny (1992) notes that, even if common belief of expected utility maximization holds at the initial node, it cannot hold after the first player passes and therefore subsequent play need not conform to backward induction. That being the case, passing at the first stage can be perfectly rational.<sup>1</sup> Because of the myriad reasons for choosing not to stop in the centipede game, it is difficult to determine why stopping at the first node is so rare empirically—is it due to a failure to reason backward or for one of these other reasons?

These demanding assumptions induced McKelvey and Palfrey (1992 p. 803) to choose the centipede game for their seminal experiment exploring alternative models since they wished to “intentionally choose an environment in which we expect Nash equilibrium to perform at its worst.” The game did not disappoint McKelvey and Palfrey, and it has consistently produced outcomes that depart radically from the predictions of Nash equilibrium (Rosemarie C. Nagel

---

<sup>1</sup>For more thorough discussions of the relationship between common knowledge of rationality and backward induction, see Aumann (1995), Elchanan Ben-Porath (1997), Asheim and Dufwenberg (2003), and Pierpaolo Battigalli and Marciano Siniscalchi (1999). Aumann (1992 p. 220) summarizes the skepticism towards backward induction in this setting eloquently, arguing that most people would say “if this is rationality, they want none of it.”

and Fang F. Tang, 1988; Fey, McKelvey, and Palfrey, 1996; Zauner, 1999; Amnon Rapoport et al., 2003; Gary Bornstein, Tamar Kugler, and Anthony Ziegelmeyer, 2004).

A notable exception to this pattern of results is the work of Ignacio Palacios-Huerta and Oscar Volij (2009), in which nearly 70 percent of their professional chess players stop the game at the first node when matched with other chess players, compared to the roughly 5% of subjects who stop in McKelvey and Palfrey (1992). Furthermore, in their artefactual field experiment every single chess Grandmaster who is given the chance to end the game on the first move does so when his opponent is known to be another chess player. When chess players are matched against students in the lab, they less frequently stop at the first node. Palacios-Huerta and Volij (2009 p. 1624, 1634) attribute the results to chess players “not satisfy(ing) even the minimal departures from common knowledge of rationality that may induce rational players to depart from backward induction” because “[b]ackward induction reasoning is second nature to expert chess players.” Further, they note that “[i]t is the rationality of a subject and his assessment of the opponent’s rationality, rather than altruism or other forms of social preferences” that is critical to determining whether backward induction will prevail.

Another strand of the experimental literature on backward induction analyzes games that attempt to untangle backward induction from assumptions about rationality and interdependent preferences (see, e.g. Binmore et al., 2002; Eric J. Johnson et al, 2002; Uri Gneezy, Aldo Rustichini, and Alexander Vostroknutov, 2007; and Dufwenberg, Ramya Sundaram, and David J. Butler, 2008).<sup>2</sup> Dufwenberg, Sundaram, and Butler (2008), and Gneezy, Rustichini, and Vostroknutov (2007) analyze zero-sum winner-take-all extensive form perfect information

---

<sup>2</sup> This research finds that assumptions about rationality and social preferences cannot fully explain departures from Nash equilibrium predictions (e.g., Binmore et al., 2002 and Johnson et al, 2002). Evidence of learning suggests that initial failures to backward induct may be due to cognitive limitations (Johnson et al, 2002; Gneezy, Rustichini, and Vostroknutov, 2007; and Dufwenberg, Sundaram, and Butler, 2008).

games with dominant strategies. Behavior in these games does not depend on social preferences or beliefs about the rationality of one's opponent. This allows for a purer measure of players' ability to recognize and implement backward induction strategies.

Our analysis brings these two strands of the literature together. Motivated by the remarkable findings of Palacios-Huerta and Volij (2009), we use professional chess players as subjects playing against other chess players, arguably giving backward induction its best chance to emerge. We conduct standard centipede games, as well as presenting results from two extremely demanding constant sum, winner-take-all "race to 100" games. As Judith D. Sally and Paul D. Sally (2003) discuss, Race to 100 is a traditional number game that involves two players who alternate choosing numbers within a given range (in our two games, either from 1 to 10 or 1 to 9). These numbers are added in sequence until one player chooses a number that makes the sum exactly equal to 100.<sup>3</sup> This player is the winner and receives a preset amount while the loser receives nothing. Similar to the games in Gneezy, Rustichini, and Vostroknutov (2007) and Dufwenberg, Sundaram, and Butler (2008), the optimal strategy in race to 100 does not depend on beliefs about other players or on distributional preferences because it is a constant sum, winner-take-all game. The dominant strategy implied by backward induction is robust to all but the most extreme types of preferences.<sup>4</sup> In contrast to Gneezy, Rustichini, and Vostroknutov

---

<sup>3</sup> The game we use is in the spirit of the games described in Sally and Sally (2003), and similar to the race game described in Gneezy, Rustichini, and Vostroknutov (2007). They denote the game by  $G(m, k)$  where players can choose any number between 1 and  $k$  and the winner is the first to make the sum equal to  $m$ . We study  $G(100, 10)$  and  $G(100, 9)$  (beginning at zero). Gneezy, Rustichini, and Vostroknutov (2007) study  $G(15, 3)$  and  $G(17, 4)$  (beginning at one). Our race to 100 games share similarities with a set of games known as "Nim" that have been analyzed in the the mathematics literature (see, e.g., Charles Bouton (1901-02); Richard Sprague, (1935-1936), and Philip Grundy (1939))

<sup>4</sup> Unless, of course, the player values the utility of her opponent more than her own utility.

(2007) and Dufwenberg, Sundaram, and Butler (2008), which require four and six steps of reasoning respectively to solve, race to 100 requires the player to reason backwards ten moves.<sup>5</sup>

In the Centipede game, our results for chess professionals are sharply at odds with those reported in Palacios-Huerta and Volij (2009) when chess players face one another in an artefactual field experiment. For instance, in our sample, chess players end the game at the first node in only 3.9 percent of the games, compared to 69 percent for Palacios-Huerta and Volij (2009) in their artefactual field experiment. And, importantly, not a single one of the sixteen Grandmasters in our experiment stop at the first node, whereas all twenty-six of the Grandmasters in Palacios-Huerta and Volij (2009) stopped at the first node. Overall, chess players in our sample behave almost exactly like standard subject pools in centipede.

Our race to 100 results suggest that failure to stop at the first node in centipede has little to do with an inability to reason backwards. In the version of race to 100 in which players choose numbers between 1 and 9, nearly 60 percent of the chess players achieve the Nash solution. Yet, among those subjects who perfectly backward induct in race to 100, not a single one stopped at the first node in centipede. Indeed, the “best inductors” in the race to 100 game had low stoppage rates in centipede at any node: passing nearly 84 percent of the time. Interestingly, this passing rate was almost identical to the rate exhibited by subjects who had no skill at backward induction in the race to 100 game. Finally, we find that small variations in the race to 100 game (the only difference being whether players choose numbers in the range 1 to 10, rather than in the range 1 to 9) influence play considerably.<sup>6</sup> This result suggests that the

---

<sup>5</sup> The problem faced by a player is somewhat easier than it might first appear, however, because the player only needs to think back on his own moves, but does not need to solve for his opponent’s optimal move. In this way, we are not testing for backward induction in the strict sense. Avinash K. Dixit (2005 p. 207) refers to this as a “rollback equilibrium.”

<sup>6</sup> Binmore et al. (2002 p. 87) similarly find that backward induction behavior is sensitive to small changes in games that are unfamiliar to players. They conclude that “backward induction would be compelling in the classical view of

ability to transfer backward induction prowess from the chess board to experimental games is quite sensitive to the particulars of the game in question.

The remainder of our paper proceeds as follows. Section 2 discusses the experimental design, including a more detailed description of the backward induction games, the subject pool, and the experimental procedure. Section 3 discusses the results. Section 4 concludes.

## II. Experimental Design

Following the bulk of the literature and Palacios-Huerta and Volij (2009), we study a version of the Centipede game that has exponentially increasing total payouts, as illustrated in Figure 1. In Figure 1, at each node, the payoffs for player I (“White”) appear in the top row and the payoffs for player II (“Black”) appear in the bottom row. Player I is the first mover. If he chooses to stop the game at the first node, player I receives \$4 and player II receives \$1. If player I chooses to continue the game at the first node, then it becomes player II’s turn to move. If player II chooses to stop the game at the second node, player I receives \$2 and player II receives \$8. If he chooses to continue the game, it becomes player I’s turn to move. The game continues until one player chooses “Stop” or the game reaches the final node. If the game reaches the final node, player I receives \$256 and player II receives \$64.

While the subgame perfect equilibrium predicts that player I choose “Stop” on his first move, few research subjects follow this strategy (Nagel and Tang, 1988; McKelvey and Palfrey, 1992; Fey, McKelvey, and Palfrey, 1996; Zauner, 1999; Rapoport et al., 2003; Bornstein, Kugler, and Ziegelmeyer, 2004).<sup>7</sup> As mentioned earlier, the literature has documented numerous reasons why players may choose to continue the game.

---

game theory, in which games are complete, literal representations of strategic interaction. But game theory is typically used not as a literal description but as a *model* of more complicated strategic interaction.”

<sup>7</sup> For a learning theory that models the influence of experience on end behavior in finite games we direct the reader to the seminal paper of Selten and Rolf Stoecker (1986). As the model suggests, such behavior is a general

A more direct test of backward induction is the race to 100 game. In this game, two players alternate choosing numbers within a given range. These numbers are added in sequence until one player chooses a number that makes the sum exactly equal to 100 (beginning from zero). This player is the winner and receives \$10, whereas the other player receives nothing. We played two variants of the race to 100 game: one in which players could choose numbers from 1-9 inclusive, and the other in which they could choose from 1-10 inclusive. In the 1-9 version, the second mover has the advantage: proper backward induction ensures victory for player II. This holds because choosing whatever number yields a sum of 10, and subsequently choosing the numbers that add to 20, 30, 40, ..., 80, 90, 100 provides a guaranteed win for player II. In the 1-10 game, however, the first mover controls her own destiny since important sums are 1, 12, 23, 34, ..., 78, 89, 100. If one player fails to backwardly induct properly on any move, the other player can guarantee victory by reaching one of those key numbers and then acting properly thereafter.

The race game's dominant strategy, constant-sum, winner-take-all reward structure eliminates concerns about loss of potential surplus, one's own social preferences, beliefs about others' social preferences, and beliefs about others' ability to backwardly induct. Consequently, play in the race games more clearly maps to tests of backward induction than do choices in the centipede game. In this manner, our approach of linking individual play across the centipede and race games is in the spirit of Binmore et al. (2002), who break backward induction into its components, subgame consistency and truncation consistency, via experimental methods.

---

phenomenon, and empirically can be found in games related to the centipede variant. For example, Johnson et al. (2002) compare two explanations of why deviations from perfect equilibrium occur in three-round bargaining games. They report that both explanations—limited cognition and social preferences—play a role. The authors propose an extensive form level-k explanation for their data. Taking this idea to centipede games, Toshiji Kawagoe and Hirokazu Tazikawa (2008) show that level-k analysis provides consistently good predictions for individual behavior.

### A. *Subject Pool*

We recruited chess players at two international open chess tournaments that took place in the spring and summer of 2008 in the US: the Chicago Open in Wheeling, IL (May 23-26) and the World Open in Philadelphia, PA (July 1-6).<sup>8</sup> While anyone entered in the tournaments was eligible to participate, we concentrated our recruiting efforts on highly ranked players.

The World Chess Federation (FIDE) and the United States Chess Federation (USCF) rank chess players using the Elo rating method.<sup>9</sup> That rating, combined with achievements in selected tournaments, qualify players for official titles. The Grandmaster (GM) title is the highest title a chess player can receive. It is followed in prestige by the International Master (IM) title; the Federation Master and the USCF Master follow – two approximately equal titles, with the former awarded by the FIDE and the latter awarded by the USCF. We categorize players who hold one or both of these titles as Masters.<sup>10</sup> Typically, GMs have an Elo rating above 2,500, IMs above 2,400, and Masters above 2,200. The experiment also included players who hold no title. We divide these players into two categories: those with a rating above 2,000 and those with a rating below 2,000. Strong club players have an Elo rating of about 1,800.

Table 1 summarizes the titles and ratings of the subjects in our study. Our sample consists of 206 chess players (103 pairs): 26 GMs, 20 IMs, 61 Masters, 46 players with no chess title rated above 2,000 and 53 players with no chess title rated below 2,000. The first movers

---

<sup>8</sup> Total prize money for the Chicago Open was \$100,000 with a top prize of \$10,000. Total prize money for the World Open was \$400,000 with a top prize of \$30,000. The average prize money payout to high rated players in these tournaments is a few hundred dollars, roughly equal to the entrance fee the tournaments charge. The hourly wage earned by players in our experiment was well above the implied hourly wage from participating in the tournament.

<sup>9</sup> Players may have both a FIDE and USCF rating. Ratings for a given player will not be identical since these ratings are based on their performance in tournaments sanctioned by either FIDE or the USCF. For more information about the rating system, see the FIDE Handbook (2008) Section B.02.10 or the USCF Handbook (2008).

<sup>10</sup> A player may hold both a Federation Master title and a USCF Master title at the same time. Players earn these titles based on their performance in tournaments sanctioned by either FIDE or the USCF.

consisted of 16 GMs, 12 IMs, 33 Masters, 18 players with no chess title rated above 2,000 and 24 players with no chess title rated below 2,000.

### *B. Experimental Procedure*

At each tournament, we rented two conference rooms in the hotels where the tournaments were held to conduct the experiments.<sup>11</sup> We ran the experiment with pairs of chess players who remained anonymous to one another. We informed each player that they would be participating in a game that requires two players who take turns in sequence, and that the other player was receiving the same instructions in another room. While participants did not know each other's identity, it is likely they assumed, given the context of the tournament, that they were paired with other chess players. Each pair played three rounds: one round of centipede and two rounds of race to 100 (one round of the 1-9 variant and one round of the 1-10 variant). We randomized whether centipede or race to 100 was played first, and likewise within the race to 100 game, whether the 1-9 or 1-10 variant was done first.<sup>12</sup> We randomly assigned each player the role of player I ("White") or player II ("Black"), and those roles were maintained throughout all three rounds.<sup>13</sup> Players did not know how many games they would play or that they would remain as either the first mover or the second mover. At the beginning of round two and round three, we informed players that they would be playing another game with the same player and that they would remain in the same mover role. The order in which the games were played does not appear to affect play in the centipede game, but does influence actions in the race to 100 game, as we discuss later.

---

<sup>11</sup> We ran the experiment while the tournament was occurring. The players generally participated between tournament rounds or during a round if their game ended early or if they took a bye. In some cases, players in contention to win the tournament requested to participate immediately after the tournament ended.

<sup>12</sup> Due to time constraints, one pair stopped early and did not play the centipede game; another pair stopped early and did not play race 1-10.

<sup>13</sup> The one exception to random assignment was that we were more likely to assign Grandmasters at the Chicago Open to the role of player I. We did this to ensure that we could observe at least one move for Grandmasters in the centipede game.

Before each game, experimenters gave players a copy of written instructions for the game and read aloud from a cue card. In order to prevent collusion, players communicated their decisions via Instant Message on computers operated by the experimenters. During each game, players recorded their own decisions and the decisions of the other player as they occurred.<sup>14</sup> After the third round, we asked players to fill out a short survey. Immediately following the experiment, we paid players their earnings from all three games privately in cash.<sup>15</sup>

The astute reader will note that we attempted to follow the field experimental design in Palacios-Huerta and Volij (2009) as closely as possible. For example, we directly informed subjects that they would be playing each game only once, as they did. This is an attempt, in the spirit of Miguel A. Costa-Gomes and Vincent Crawford (2006), to study strategic thinking in an environment without learning. And, at no time did we mention other games that subjects might be playing later in the experiment, thus the game played first is the strategic analog to a one-shot game.

One difference between our design and the field design of Palacios-Huerta and Volij (2009) is that our subjects play both the centipede game and the race to 100 games with the same partner. The fact that we have some subjects play the centipede game first while other subjects play the centipede game after the race games permits a test of learning across domains. In practice, we find that learning is minimal (i.e., ordering does not matter) across the centipede game and the race to 100 games.

We should be clear that while our centipede game represents a direct replication of the artefactual field experiment in Palacios-Huerta and Volij (2009), in no way are we trying to

---

<sup>14</sup> In the race to 100 games, subjects recorded the number chosen and current sum at the end of each turn. The experimenters confirmed that players had correctly recorded these numbers before proceeding with play.

<sup>15</sup> See the online Appendix for a copy of the written instructions, the cue cards read aloud by experimenters, and the survey.

replicate their lab experiment. The latter used a random re-matching design, had each subject play ten centipede games, and varied opponent type—chess players versus other chess players, chess players versus students, and students versus students. Such an approach permits an analysis of questions beyond those of direct import herein.

### III. Experimental Results

Table 2 and Figure 2 present results on the probability of stopping at each node in the centipede game, conditional upon reaching that node. The top row in Table 2 pools data across players; the remaining rows parse the data according to the chess ranking of the player making a decision at the node. Each column in the table corresponds to a different node. Looking first at the results pooling all players, we find that stop probabilities are low at early nodes and generally increase over the course of the game. For example, in only 3.9 percent of the 102 pairs does the player stop the game at the first node.

Figure 2 reveals that roughly 10 percent of the games end at the second node. Nearly 45 percent of the games make it to at least node 5 (at which point player 1 gets \$64 and player 2 receives \$16). Remarkably, 37 percent of players who reach the sixth node choose to continue the game to the final node (for a sure loss of \$64).<sup>16</sup> This result is consistent with many underlying motivations, including positive reciprocity—a player who has been generous at earlier nodes is rewarded by the opponent at the final node, even though the opponent suffers a substantial financial loss in doing so—more general social preferences, and bounded rationality.

Given the high pass rates observed in the data, passing maximizes expected returns at every node except the last one (where passing guarantees a loss). For instance, in our data the average final payoff to player 1 when he elected to pass at the first node was \$44.85, compared

---

<sup>16</sup> Among those who played the centipede game before either of the race to 100 games, 1.9 percent of the pairs stopped at the first node, 7.7 percent at the second node, and 77 percent made it to at least node 5. 7.7 percent of the pairs reached the final node.

to a guaranteed \$4 from stopping at the first node (average earnings in total were \$38.14). These results are consistent with previous studies of the centipede game using student subjects.

In terms of level of chess expertise, it is interesting to note that in sixteen pairs with a Grandmaster as player 1, not once did the Grandmaster stop at the first node. Only one of nine Grandmasters who had the chance ended the game at the second node, and in only one of fifteen opportunities did a grandmaster end the game at the third node. The same general pattern is true among International Masters, as well as lower ranked players. Overall, the centipede results for our chess sample look very much like the empirical results with standard subject pools and strikingly different than the earlier findings on chess players reported in Palacios-Huerta and Volij (2009).

#### *A. Race to 100 games*

Tables 3 and 4 summarize empirical results for the race to 100 games. Table 3 splits the data according to the first point at which the game is “solved,” where solved means that one player gets onto a number that guarantees victory and plays optimally at every move from that point onward. The top two rows in Table 3 show results for the variant of the game in which players choose 1 to 9; the bottom two rows correspond to the 1 to 10 treatment. In each case we report both the “key numbers” that ensure victory and the percentage of games that are first solved at that particular “key number.”

The top portion of Table 3 demonstrates that in the 1 to 9 treatment, 57.3 percent of pairs solved the game as early as possible (i.e., at the number 10). By the number twenty, roughly two-thirds of the pairs have solved the game. If the solution is not achieved by then, the game is likely not solved until near the end. Interestingly, the chess players do much worse in the 1 to 10 treatment: only 12.6 percent of these cases are solved on the first move (which requires the

player to choose 1) and in only roughly 20 percent of cases is there a solution by the second key number (which is 12).<sup>17</sup> Remarkably, nearly two-thirds of the 1 to 10 treatment are not solved until number 78 or higher.

We find it striking that even among a subject pool that has extensive experience with backward induction, the seemingly minor change of shifting the “key numbers” from numbers ending with zero leads to a sharp reduction in success in solving the problem.<sup>18</sup> This result is consistent with the power of subtle changes reported in many psychology experiments as well as Binmore et al. (2002 p. 87), who find that backward induction behavior of players unfamiliar with the game is quite sensitive to minute changes in the game, and also with the findings of Adriaan de Groot (1965) regarding the difficulty chess players have in generalizing their skills in unfamiliar settings, even within relatively narrow contexts.

Tables 4a and 4b present empirical results of the race to 100 game in terms of implied probabilities of solving the problem, conditional on reaching that point without a solution previously being obtained. Table 4a (4b) shows the 1 to 9 (1 to 10) treatment. The first column of Table 4a corresponds to the first chance a player has to solve the problem (i.e., in the 1 to 9 variant, the first action by player 2, or the second move by player 1 if player 2 fails to solve the problem on the first move). The other columns map to each of the relevant “key numbers.” The top row of the table provides pooled results, whereas subsequent rows split the data by chess ranking. The number of opportunities that arise at each node is presented in parentheses.

---

<sup>17</sup> These results are consistent with Dufwenberg, Sundaram, and Butler (2008) who find that 14 percent of players solve a related game on the first move in the first round of play. Gneezy, Rustichini, and Vostroknutov (2007) do not separately report first round results for the race game they analyze. Yet, pooling the first five rounds of play, they find 46.5 percent of players make no errors on the first move. We should note that, similar to the centipede game results discussed above, subjects who play the race games before the centipede game behave similarly to those playing the race games after the centipede game.

<sup>18</sup> The null hypothesis of equal probabilities of finding an early solution to the 1 to 9 and 1 to 10 treatments is strongly rejected by the data.

Overall, in the 1 to 9 treatment, 39 percent of players solve the game on their first move. This number is lower than the 57.3 percent of *games* that are solved by the number 10 because in cases where the game is not solved by one player at the first chance, the other player often also fails to successfully solve it.<sup>19</sup> If the game is not solved early, then the probability that it is solved at a middle node is less than 10 percent. By the number 80, the hazard rate for solving it rises to 24 percent, and by 90, the solution rate is nearly three-quarters. Grandmasters are the group most likely to solve the game on their first chance (50 percent likelihood), but among the Grandmasters who fail to solve the game immediately, their performance is weak. Generally, we do not find stark differences in performance across chess rankings.

Data trajectories in the 1 to 10 game are similar, except that the solution rates early in the game are much lower than in the 1 to 9 treatment. Once again, if a solution is not reached near the beginning, it is unlikely that the game will be solved until near the end: about 1 in 5 players who reach the 78 node solve it there, and about three-fourths of the players get the right answer at 89. Grandmasters do somewhat better than other chess players with respect to finding an early solution to the 1 to 10 game, but a test of the null hypothesis that Grandmasters have the same probability of solving this game at the first opportunity is rejected only at the  $p < 0.10$  level.

To explore whether learning takes place within the course of the experiment, we examine whether the proportion of games solved at each stage of the race to 100 depends upon whether this is the first or second race to 100 game for the players (recall that every subject plays both the 1 to 9 and the 1 to 10 version). As shown in the bottom two rows of Table 5, when the 1 to 10 game is played after the 1 to 9 game, performance in the 1 to 10 game is improved. A solution is reached by the key number 12 nearly 30 percent of the time if players first see the 1 to 9 game,

---

<sup>19</sup> Note also that it is possible for neither player to solve the game on the first try, but for the game to be solved at 10 if both players' first actions are low numbers, and then the third number selected solves the problem. This happens in one instance in the data.

compared to roughly half that rate if they have not. This difference is statistically significant at the  $p < 0.01$  level. Interestingly, the top two rows of the table demonstrate that when the 1 to 9 game is played second, performance on the 1 to 9 game actually *diminishes* slightly relative to when the 1 to 9 game is played first, with that difference statistically significant at the 0.05 level. In contrast to the 1 to 9 game, in the 1 to 10 variant the solution frequently comes very late in the game, and players do not appear to generalize from the fact that the game can be solved by backward induction once the numbers are close to 100 to the broader lesson that this class of games is susceptible to backward induction from the very first move.

One explanation for these differences across games is that the pattern that emerges from the backward induction approach in the 1 to 9 game is easily recognizable: 60, 70, 80, 90, 100. To a player who is not using backward induction initially, the fact that the opponent always chooses a number that ends on zero and wins using that strategy sends a strong cue as to the nature of the solution to the game. In contrast, the pattern of choices in the 1 to 10 game is much less obvious: 56, 67, 78, 89, 100. A player could ride the backward induction strategy to victory without the opponent even recognizing that there was a pattern to the first player's actions.<sup>20</sup>

Interestingly, we also find some evidence that playing the centipede game first interferes with solving the 1 to 10 game. If race to 100 is played first, 20.2% of pairs solve the 1 to 10 game on the first move (i.e., by choosing 1). This percentage drops to 5.88% when centipede is played first. Similarly, the percentage of games solved on the final node (i.e., 89) increases from 33.3 percent to 55.9 percent when race to 100 is played first and second, respectively. The effect

---

<sup>20</sup> Using a related game, Dufwenberg, Sundaram, and Butler (2008) find that playing an easier variant before a harder one increases the probability that players will first recognize that there is an analytic solution and second will solve for the dominant strategy (there, the easier variant required fewer steps of reasoning than the harder variant). However, in contrast to our results, they do not find that playing the harder variant first hurts play in the easier variant. This may be because the easier variant only requires two steps of reasoning. Gneezy, Rustichini, and Vostroknutov (2007) also find that learning takes place over the course of repeated race games, with subjects switching from forward induction to backward induction as experience increases. Although their subjects played race games of varying difficulty, they do not discuss the effect of game order on performance.

on the 1 to 9 game is similar in direction but the differences are small. One explanation for this result is that the players see centipede as a game of cooperation as opposed to a backward induction game, which interferes with their ability to conceptualize the race to 100 game as a backward induction game.<sup>21</sup>

*B. Is skill at backward induction correlated with stopping early in the centipede game?*

As noted earlier, there are many reasons why a player in the centipede game might not choose to stop the game. Inability to backward induct is one of those explanations.<sup>22</sup> Using performance in the race to 100 game as a measure of skill at backward induction, we are able to test whether those who successfully backward induct in the race to 100 game are more likely to stop at a given node in the centipede game.

There are 15 players in our sample who backward inducted perfectly in the race to 100 games, i.e. they made the optimal backward induction strategy every chance they were given in both games. The top row of Table 6 reports the probability that these 15 players stop at each node in the centipede game, conditional on reaching that node. The number of times each node is reached is shown in parentheses. Despite the fact that these 15 players proved themselves adept at backward induction in the race to 100 games, not once out of ten opportunities did they choose to stop at the first node in the centipede game. At nodes two through four, these players never stopped the game more than 25 percent of the time. At node five, they stopped two out of three times. Tellingly, on the two occasions in which they reached node six, these players elected to pass both times, even though no backward induction whatsoever is required to see that passing at the last node lowers one's own payoff, suggesting that other forces, such as social

---

<sup>21</sup> Another explanation is subject fatigue.

<sup>22</sup> Because the subjects play each game variant only once, our measure of backward induction ability captures players' ability to backward induct in response to a novel situation rather than their ability to learn the backward induction strategy over the course of repeated play.

preferences, are at work. Despite demonstrated ability to backward induct flawlessly in the race to 100 games, this group of players elected to stop the centipede game in only 17 percent of the opportunities that they faced. These results argue against interpreting failure to stop in the centipede game as evidence of an inability of an individual to backward induct.

The second row of Table 6 shows results for players who exhibited some backward induction ability in the race games, although they did not play those games perfectly. These “Second-best” inductors are players who solved one race to 100 game on their first move or solved both race to 100 games before the last node, but made at least one mistake in those games. Sixty-six subjects fit this classification. Similar to the best inductors, these players rarely stopped the centipede game at early nodes. Not once in 17 opportunities did they stop at node 1 when given the opportunity, and in only 10.6 percent of the cases did they stop at the second node.<sup>23</sup> These players were more likely than the perfect inductors to stop centipede at other nodes, and overall stopped the centipede game in 31 percent of the chances they were given.

Our third classification of players includes those who did not qualify for the top two categories for backward induction, but did solve at least one of the race to 100 games prior to the last node. Thirty six players fell into this category. These subjects play centipede much like the second best inductors, with an overall stopping rate of 32 percent.

The last group of players includes those who did not solve either of the race to 100 games prior to the final node. Nearly forty percent of our subjects fall into this category. Interestingly, this set of players, who showed no proficiency for backward induction, played the centipede game most like the perfect backward inductors, passing at very high rates. In sum, we find no

---

<sup>23</sup> These second best inductors have many more opportunities at the second node than the first node because the most common way to qualify was to solve correctly the 1 to 9 version of the race to 100 game on the first move. In order to have that chance, one had to be a second mover in that game, and if you were a second mover in one game, you moved second in all games.

evidence that stopping in the centipede game is systematically related to backward induction performance in the race to 100 games, calling into question the validity of using centipede games to draw inferences about backward induction.

#### **IV. Conclusion**

In this study, we explore the behavior of world class chess players in complementary games that lend insights into backward induction prowess. We find that these players exhibit substantial abilities to backward induct in games appropriate for tests of backward induction, but do not choose the backward induction solution in the centipede game. This behavior cannot easily be attributed to an inability to backward induct since it is uncorrelated with demonstrated backward induction ability in the more appropriate tests of backward induction.

Indeed, given the actual play of opponents, such cooperative behavior in the centipede game is wealth maximizing. One explanation for this high degree of cooperation in our experiment is that cooperative arrangements are common in tournament chess. For example, anecdotal evidence suggests that it is common for chess players to agree to a draw (tie) prior to a game towards the end of tournaments when such collusive behavior is jointly beneficial. Chess players also report agreeing in advance of games to 60-40 splits (60% of tournament payoffs to the game winner, 40% to the loser) in order to reduce the variance of payoffs.<sup>24</sup>

---

<sup>24</sup> Consistent with such behavior, in our own experiment one Grandmaster, aware that his opponent was likely another chess player at this tournament, offered the following unsolicited opinion to describe his rationale in the centipede game: "If this man were my enemy and he would use the money to hurt me, then I would stop. But this man is my friend. I can trust him. I will continue." He reached the final node.

## References

- Asheim, Geir B., and Martin Dufwenberg. 2003. "Deductive Reasoning in Extensive Games." *The Economic Journal*, 113(April): 305-325.
- Aumann, Robert J. 1988. "Preliminary Notes on Integrating Irrationality into Game Theory." International Conference on Economic Theories of Politics, Haifa.
- Aumann, Robert J. 1992. "Irrationality in Game Theory." In *Economic Analysis of Markets and Games: Essays in Honor of Frank Hahn*, ed. Frank Hahn and Partha Dasgupta, 214-227. Cambridge, MA: MIT Press.
- Aumann, Robert J. 1995. "Backward Induction and Common Knowledge of Rationality." *Games and Economic Behavior*, 8(1): 6-19.
- Battigalli, Pierpaolo, and Siniscalchi Marciano. 1999. "Hierarchies of Conditional Beliefs and Interactive Epistemology in Dynamic Games." *Journal of Economic Theory*, 88(1): 188-230.
- Ben-Porath, Elchanan. 1997. "Rationality, Nash Equilibrium and Backwards Induction in Perfect-Information Games." *Review of Economic Studies*, 64(1): 23-46.
- Binmore, Kenneth G. 1987. "Modeling Rational Players." *Economics and Philosophy*, 3: 179-214.
- Binmore, Kenneth G., John McCarthy, Giovanni Ponti, Larry Samuelson, and Avner Shaked. 2002. "A Backward Induction Experiment." *Journal of Economic Theory*, 104(1): 48-88.
- Bornstein, Gary, Tamar Kugler, and Anthony Ziegelmeyer. 2004. "Individual and Group Decisions in the Centipede Game: Are 'Groups' More Rational Players?" *Journal of Experimental Social Psychology*, 40(5): 599-605.
- Bouton, Charles L. (1901-1902). "Nim, A Game with a Complete Mathematical Theory," *Annals of Mathematics*, 3(1/4): 35-39, 2<sup>nd</sup> Ser.
- Costas-Gomes, Miguel A., and Vincent P. Crawford. 2006. "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study." *American Economic Review*, 96(5): 1737-68.
- De Groot, Adriaan. 1965, *Thought and Choice in Chess*. The Hague: Mouton.
- Dixit, Avinash K. 2005. "Restoring Fun to Game Theory." *Journal of Economic Education*, 36(3): 205-219.

- Dufwenberg, Martin, Ramya Sundaram, and David J. Butler. 2008. "Epiphany in the Game of 21." <http://www.u.arizona.edu/~martind1/>.
- Fey, Mark, Richard D. McKelvey and Thomas R. Palfrey. 1996. "An Experimental Study of Constant-Sum Centipede Games." *International Journal of Game Theory*, 25(1): 269-287.
- FIDE Handbook. 2008. World Chess Federation. <http://www.fide.com/component/handbook/>.
- Gneezy, Uri, Aldo Rustichini, and Alexander Vostroknutov. 2007. "I'll Cross That Bridge When I Come to It." <http://www.vostroknutov.com/pdfs/racebeh.pdf>.
- Grundy, Philip. (1939). "Mathematics and Games," *Eureka*, 2: 6-8.
- Johnson, Eric J., Colin Camerer, Sankar Sen, and Talia Rymon. 2002. "Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining." *Journal of Economic Theory*, 104(1): 16-47.
- Kawagoe, Toshiji, and Hirokazu Takizawa. 2008. "Level-K Analysis of Experimental Centipede Games." [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1289514](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1289514).
- Kreps, David M. 1990. *A Course in Microeconomic Theory*. New Jersey: Princeton University Press.
- McKelvey, Richard D., and Thomas R. Palfrey. 1992. "An Experimental Study of the Centipede Game." *Econometrica*, 60(4): 803-836.
- Nagel, Rosemarie C., and Fang F. Tang. 1988. "Experimental Results on the Centipede Game in Normal Form: An Investigation on Learning." *Journal of Mathematical Psychology*, 42(2-3): 356-384.
- Palacios-Huerta, Ignacio, and Oscar Volij. 2009. "Field Centipedes." *American Economic Review*, 99(4): 1619-1635.
- Rapoport, Amnon, William E. Stein, James E. Parco and Thomas E. Nicholas. 2003. "Equilibrium Play and Adaptive Learning in a Three-Person Centipede Game." *Games and Economic Behavior*, 43(2): 239-265.
- Reny, Philip J. 1988. "Common Knowledge and Games with Perfect Information." *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 2: 363-369.
- Reny, Philip J. 1992. "Rationality in Extensive Form Games." *Journal of Economic Perspectives* 6 (Autumn): 103-118.
- Rosenthal, Robert W. 1981. "Games of Perfect Information, Predatory Pricing and Chain Store Paradox." *Journal of Economic Theory*, 25(1): 92-100.

Sally, Judith D., and Paul J. Sally. 2003. *TriMathalon: A Workout Beyond the School Curriculum*. Natick, MA: AK Peters, Ltd.

Selten, Reinhard. 1965. "Spieltheoretische Behandlung eines Oligopolmodells mit Nachfragertragheit," *Zeitschrift für die Gesamte Staatswissenschaft*, 12: 301-324.

Selten, Reinhard and Rolf Stoecker. 1986. "End Behavior in Sequences of Finite Prisoner's Dilemma Supergames." *Journal of Economic Behavior and Organization*, 7(1): 47-70.

Sprague, Richard (1935-1936). "Über Matematische Kampfspiele," *Tohoku Mathematics Journal*, 41(13): 438-444.

USCF Handbook. 2008. United States Chess Federation. <http://main.uschess.org/index.php>.

Zauner, Klaus G. 1999. "A Payoff Uncertainty Explanation of Results in Experimental Centipede Games." *Games and Economic Behavior*, 26(1): 157-185.

## Figure 1: Centipede Game Payoff Structure

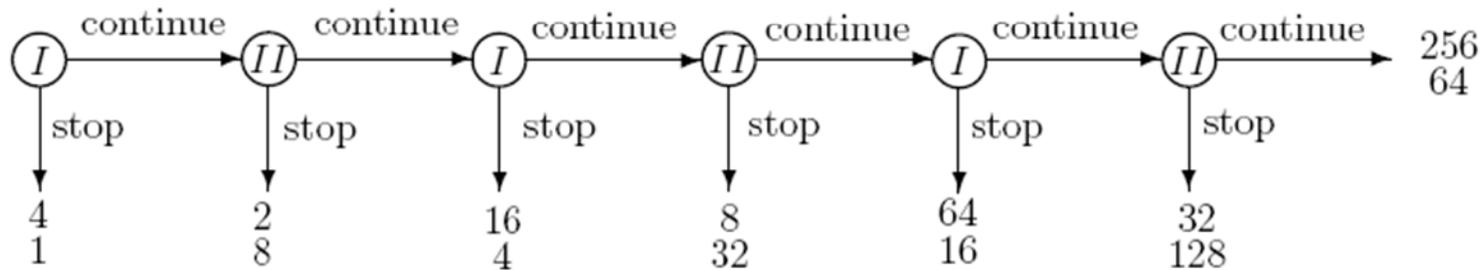
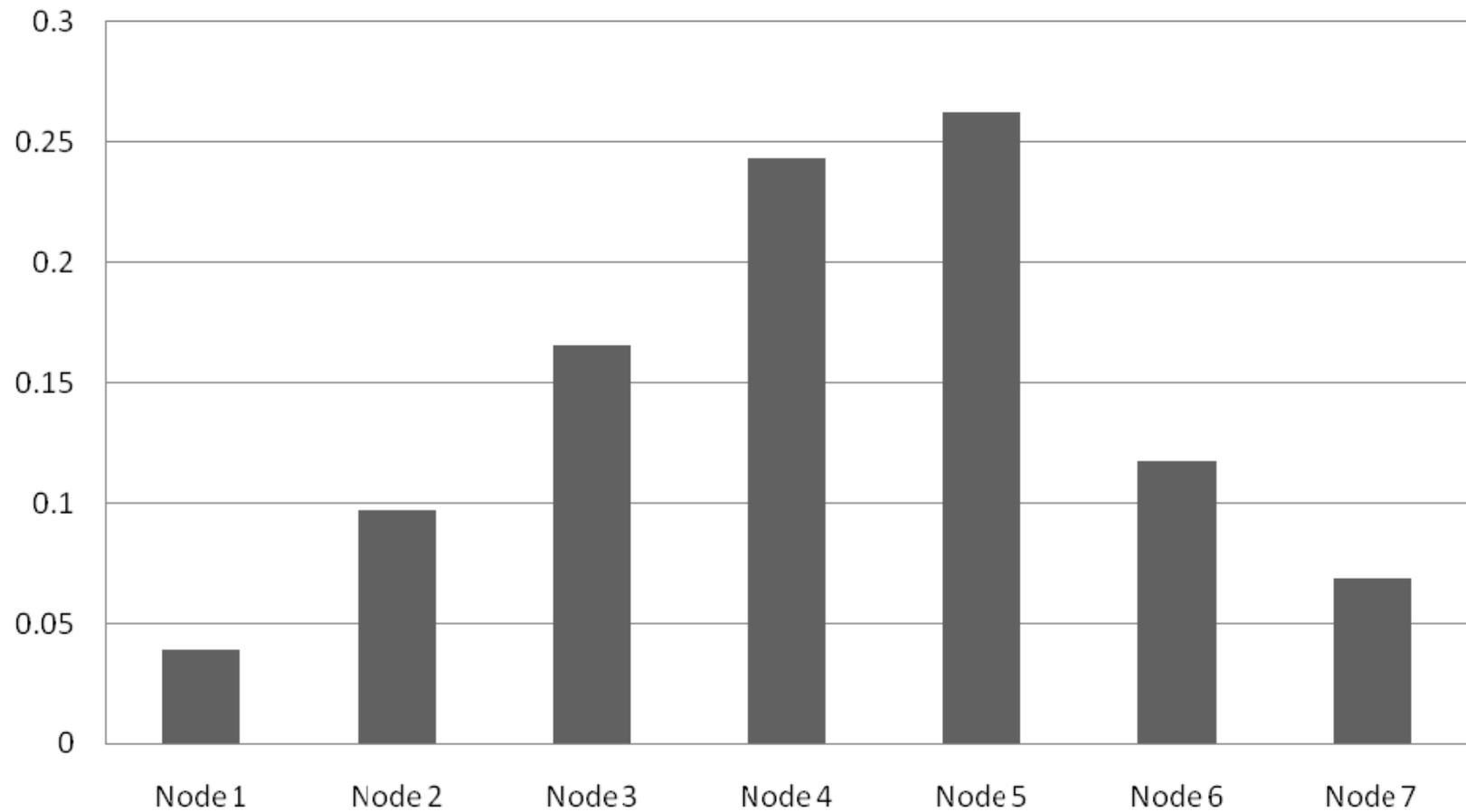


Figure 1 displays the centipede game payoff structure. A stop in each round results in player I receiving the top payoff, and player II receiving the bottom. The player whose turn it is marked in the circle.

<b>Table 1: Summary of Participants</b>			
	USCF Rating	FIDE Rating	Number of Players
Grandmasters	2472-2763	2265-2637	26
International Masters	2273-2538	2227-2488	20
FIDE & USCF Masters	2009-2497	2037-2497	61
Other Chess Players > 2000	2000-2610	2026-2531	46
Other Chess Players < 2000	852-2064	1789-2341	53
Total Number	852-2763	1789-2637	206

Table 1 reports subject pool results. Columns correspond to title, the range of USCF ratings, FIDE ratings, and subject counts.

**Figure 2: Distribution of Centipede Game Stopping Nodes**



<b>Table 2: Summary of Centipede Results- Implied Stop Probability</b>						
	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6
All Chess Players	0.039 (102)	0.102 (98)	0.193 (88)	0.352 (71)	0.587 (46)	0.632 (19)
Grandmasters	0 (16)	0.111 (9)	0.067 (15)	0 (7)	0.636 (11)	1 (2)
International Masters	0 (12)	0 (8)	0.083 (12)	0.625 (8)	0.625 (8)	0 (1)
Masters	0.063 (32)	0.077 (26)	0.259 (27)	0.318 (22)	0.455 (11)	0.5 (8)
>2000	0.056 (18)	0.154 (26)	0.154 (13)	0.2 (15)	0.5 (6)	1 (2)
<2000	0.042 (24)	0.103 (29)	0.286 (21)	0.526 (19)	0.7 (10)	0.667 (6)

Table 2 reports the distribution of implied stop probabilities for players in the centipede game. Columns correspond to the conditional probability that a player will stop at that node, given the chance to do so. Odd nodes refer to player one's choices, while even nodes refer to player two's choices. Number of opportunities observed is displayed in parentheses below.

<b>Table 3: Summary of Race to 100 Results</b>									
Node at which Game Solved	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>
Number at which Game Solved (1-9)	10	20	30	40	50	60	70	80	90
Percentage of Time Solved (1-9)	0.573	0.087	0.029	0.039	0.019	0.01	0.029	0.078	0.136
Number at which Game Solved (1-10)	1	12	23	34	45	56	67	78	89
Percentage of Time Solved (1-10)	0.126	0.087	0.019	0.01	0.01	0.029	0.049	0.214	0.447

Table 3 reports the distribution of nodes at which a race to 100 game was solved. Rows 1 and 3 report the "key number" from which a win may be forced. Rows 2 and 4 report the percent of the time a corresponding game was solved at that node; a solution is a choice of number that summed to a "magic number", in conjunction with never deviating from subsequent "key numbers" afterward.

Table 4A: Implied Probabilities 1-9										
	First	10	20	30	40	50	60	70	80	90
All	0.392 (158)	0.333 (3)	0.147 (34)	0.061 (49)	0.083 (48)	0.048 (42)	0.026 (38)	0.086 (35)	0.242 (33)	0.737 (19)
GM	0.5 (16)	. (0)	0 (4)	0 (4)	0.167 (6)	0 (5)	0 (5)	0 (4)	0 (5)	0.5 (2)
IM	0.333 (12)	. (0)	0 (4)	0.333 (3)	0 (4)	0 (4)	0 (2)	0 (2)	0.75 (4)	. (0)
Master	0.422 (45)	. (0)	0.4 (5)	0 (10)	0.077 (13)	0.083 (12)	0 (9)	0.111 (9)	0.222 (9)	1 (5)
>2000	0.357 (42)	0.333 (3)	0 (13)	0.118 (17)	0.167 (12)	0.091 (11)	0.091 (11)	0 (8)	0 (6)	0.6 (5)
<2000	0.372 (43)	. (0)	0.375 (8)	0 (15)	0 (13)	0 (10)	0 (11)	0.167 (12)	0.333 (9)	0.714 (7)

Table 4A reports the distribution of implied probabilities for race to 100 (1-9). Columns correspond to the "key number" from which a win can be forced. Columns correspond to the conditional probability that a player will solve a game at that number, given the chance to do so. Number of opportunities observed is displayed in parentheses below.

<b>Table 4B: Implied Probabilities 1-10</b>									
	First	12	23	34	45	56	67	78	89
All	0.098 (193)	0.095 (21)	0.022 (91)	0.008 (120)	0.008 (125)	0.026 (116)	0.047 (107)	0.198 (111)	0.754 (61)
GM	0.2 (25)	0 (2)	0 (5)	0 (16)	0.071 (14)	0.077 (13)	0.1 (10)	0.286 (14)	0.75 (4)
IM	0.105 (19)	0.333 (3)	0 (9)	0 (11)	0 (10)	0.091 (11)	0.083 (12)	0.167 (12)	1 (5)
Master	0.107 (56)	0 (5)	0.043 (23)	0 (26)	0 (36)	0 (31)	0.04 (25)	0.226 (31)	0.8 (20)
>2000	0.024 (42)	0.167 (6)	0 (21)	0 (36)	0 (36)	0.034 (29)	0.103 (29)	0.045 (22)	0.647 (17)
<2000	0.098 (51)	0 (5)	0 (28)	0.031 (32)	0 (28)	0 (32)	0 (33)	0.219 (32)	0.75 (16)

Table 4B reports the distribution of implied probabilities for race to 100 (1-10). Columns correspond to the "key number" from which a win can be forced. Columns correspond to the conditional probability that a player will solve a game at that number, given the chance to do so. Number of opportunities observed is displayed in parentheses below.

<b>Table 5: Percentage of Games Solved at Each Node</b>									
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>
Race to 100 1-9	10	20	30	40	50	60	70	80	90
1-9 Played First (N=52)	0.615	0.115	0.038	0.019	0.019	0	0.019	0.058	0.115
1-9 Played Second (N=51)	0.529	0.059	0.02	0.059	0.02	0.02	0.039	0.098	0.157
Race to 100 1-10	1	12	23	34	45	56	67	78	89
1-10 Played First (N=51)	0.078	0.059	0	0.02	0	0.059	0.039	0.176	0.569
1-10 Played Second (N=51)	0.176	0.118	0.039	0	0.02	0	0.059	0.255	0.333

Table 5 reports the distribution of when a race was solved, conditional on the order races were played in. Columns correspond to the "key number" at which a game was solved. One pair of players dropped out after playing the race (1-9) first, which accounts for the differing number of observations.

<b>Table 6: Centipede Behavior by Induction Ability: Implied Stop Probabilities</b>							
	N	F1	F2	F3	F4	F5	F6
Best	15	0 (10)	0.2 (5)	0.125 (8)	0.25 (4)	0.667 (3)	0 (2)
Second Best	66	0 (17)	0.106 (47)	0.375 (16)	0.412 (34)	0.571 (7)	0.889 (9)
Second Worst	36	0.1 (20)	0.133 (15)	0.278 (18)	0.417 (12)	0.833 (12)	1 (1)
Bad	87	0.036 (55)	0.065 (31)	0.109 (46)	0.238 (21)	0.458 (24)	0.429 (7)

Table 6 displays implied stop probability by inductor ability, rather than title for players. Odd numbered columns refer to player one's decisions, even numbered columns refer to player two's decisions.