# How Does AI Improve Human Decision-Making? Evidence from the Release of the AI-Powered Go Program

Sukwoong Choi[a] Namil Kim[b,*] Junsik Kim[c] Hyo Kang[d]

May 2021

**Abstract**

How does AI improve human decision-making? Answering this question is challenging because it is difficult to assess the performance of each decision and to disentangle AI's influence on decisions. We study professional Go games, which provide a unique opportunity to overcome such challenges. In 2016 the AI-powered Go program (APG) unexpectedly beat the best human player, surpassing the best human knowledge and skills accumulated over thousands of years. To investigate the impact of APG, we compare human moves to AI's superior solution set, before and after the initial public release of APG. Our analysis of 750,990 moves in 25,033 games by 1,242 players reveals that APG significantly improved professional players' probability of winning for each move. This improvement is most prominent in the early stage of a game when uncertainty is higher and is driven mainly by reducing the likelihood of human *errors* or a *critical mistake* during the game. Furthermore, junior players—who are more open to and better able to utilize APG—benefit more from APG than senior players do, suggesting generational inequality in AI adoption and utilization.

**Keywords:** Artificial Intelligence (AI), Technology adoption, Decision-making, Human capital, Professional Go players, AI adoption inequality.

[a] MIT Sloan School of Management, MIT Initiative on the Digital Economy.
[b,*] School of Management, Harbin Institute of Technology. Corresponding author. namil.kim@hit.edu.cn.
[c] Electrical Engineering, KAIST.
[d] Marshall School of Business, University of Southern California.

## 1. INTRODUCTION

Artificial Intelligence (AI) is the newest general-purpose technology (GPT) in the foreseeable future (Trajtenberg 2018, Goldfarb et al., 2020) and has the potential to affect all aspects of the economy—including the way we innovate, do business, and organize (Cockburn et al. 2019). From the perspective of the economics of digitization (Goldfarb and Tucker 2019), AI will bring huge changes to our economy and to society. This is because AI can provide humans with high-quality predictions at relatively low cost, leading to a sharp decline in prediction cost (Agrawal et al. 2018). AI has already outperformed human professionals in certain domains—such as strategic games (Silver et al., 2017), medical diagnosis (Wang et al, 2019), and financial analysis (Cao et al., 2021). Furthermore, the domains where AI outperforms humans are rapidly expanding (Agrawal et al. 2018). This phenomenon raises an interesting and timely question about how AI affects humans in these domains.

Recent studies have focused on the effect of AI on job replacement (Acemoglu et al., 2020, Luo et al., 2019). However, AI may have broader effects beyond changes in employment. As a GPT, AI is likely to lead to substantial change in human decision-making and performance. Several empirical studies have examined AI's effect on human decision-making at different skill levels and in different jobs (Cao et al., 2021, Cowgill 2020, Kleinberg et al. 2017, Wang et al., 2020). Empirical studies in this area are challenging because it is difficult to (1) observe each decision by humans and assess the performance and (2) disentangle AI's influence on such decisions. In addition, given that the development of AI has happened recently, data availability has been limited in terms of its quantity and time span. This study aims to overcome these challenges; it uses granular data to investigate the impact of AI on professionals' decision-making and the mechanisms at the single-choice level.

We study how AI enables human professionals to make better choices by improving their heuristics or everyday practices in decision-making. We ask (1) whether AI improves human decision-making and its performance, (2) what are the mechanisms by which this occurs, and (3) what are the heterogeneous effects of access to and attitudes toward AI. Our empirical context is the professional Go market. Professional Go games provide a unique opportunity to study the impact of AI and to overcome the challenges of observing human decisions and untangling AI's influence. Each human decision (a move in the game) is observable and can be assessed precisely in terms of whether it increases or decreases the overall probability of winning and by what margin. Human players have accumulated knowledge and skills in Go games over thousands of years, but the introduction of the AI-powered Go program (henceforth, APG) suddenly changed how professional Go players learn and play the game. The historic Go match (AlphaGo vs. Sedol Lee) was held in 2016; in this game, AI beat the best human player for the first time. The first open APG, Leela, became available to players in February 2017. Quantitative assessments (as shown in Figure 2) and secondary qualitative data indicate that professional Go players have used APGs heavily in their training. The great advantage of this context is that it allows us not only to observe the changes in strategic choices of professional Go players before and after APG but also to measure the performance of Go players' decisions and to compare them with the performance of AI.

We calculated the changes in the probability of winning for every single move (i.e., 750,990 decisions) in 25,033 major games held from 2015 through 2019 and involving 1,242 professional Go players. The results show that the quality of moves by professional Go players improved substantially following the release of APG. Before APG, the winning probability of each move by professional Go players averaged 2.47 percentage points lower than the moves of APG.

This gap decreased by about 0.756 percentage points (or 30.5 percent) after the release of APG. This improvement in the quality of moves eventually leads to the final win of the game. Interestingly, this improvement is most prominent in the early stage of a game when higher uncertainty is exhibited.

Further analyses indicate that the improvements in the quality of moves are driven mainly by reducing the number of human *errors* (moves where the winning probability drops by 10 percentage points or more) and the value of the *most critical mistake* (the biggest drop in winning probability during the game). The number of errors per game decreased by 0.15–0.50 and the likelihood of a *critical mistake* decreased by 4–7 percentage points. To test for the heterogeneity in AI impact, we examine whether junior and senior players exhibit differential effects. We find that the improvement by APG is more prominent for junior professional players who are open to and capable of utilizing APGs than for senior professional players.

The findings from the professional Go market may not be readily applicable to other contexts; the application of the study's findings requires careful assessment of the context. Yet, the lessons from this study can still provide important implications for decision-making processes by business managers and policy makers. Studies highlight that the decisions made in games are similar to those made by managers (and policy makers) in that they use similar intuitive techniques under uncertainty and time constraints (e.g., Bechara et al. 1997; Mintzberg 1987, 1994; Miric et al. 2020; Simon 1987; Simon and Chase 1988).

To the best of our knowledge, this study provides the most comprehensive evidence of the effect of AI on human decision-making processes and its performance at the level of individual decisions. The wide range of outcomes we investigate—quality of moves, errors, mistakes, and winning—suggests that AI has a substantive effect on professionals' decisions and performance

by providing frontier knowledge and new strategies that previous GPT could not provide. Our context also provides a deeper understanding of how AI affects every single human decision and the corresponding outcomes that are difficult to measure.

The improvements made by APG make us realize not only how even the best human professionals have heuristics in decision-making but also how can AI support them in making better decisions. In particular, AI can improve human decision-making in a short-paced, uncertain environment; the power of this improvement is significant to the extent that AI could overturn human knowledge and wisdom accumulated over thousands of years. Further, the fact that junior professionals benefited more from the AI-powered program has important implications for AI adoption, utilization, and inequality.

The paper proceeds as follows. Section 2 discusses the related literature on human decision-making, and the roles of AI and age. Section 3 explains the data, research design, and identification strategy used for empirical analysis. Section 4 presents the results and robustness checks. Section 5 provides further analyses. The discussion and conclusion are provided in Section 6.

## 2. AI AND DECISION-MAKING

### 2.1. The role of AI in human decision-making

When making decisions, as humans we tend to draw on our conceptualization of the future as input into the decision-making process (Lindebaum et al. 2020; Mintzberg 1987, 1994). We also depend on knowledge of causality, which we actively develop to understand how past actions impact future outcomes. Through these processes, we can judge and learn from situations—even unexpected situations—to improve our decision-making processes and outcomes (Lindebaum et

al. 2020; Mintzberg 1994). However, individuals are limited in their ability to process information, which slows down learning and limits its scope (Cyert and March 1963; Galbraith 1974; Simon 1955, 1958). We do not always optimize decision-making (Kalberg 1980). For instance, our choices are often affected by rigidity to change and other routines, which lead to learning myopia (Levinthal and March 1993).

Acknowledging these limitations, researchers have studied incompleteness in managerial decision-making (Dane and Pratt 2007; Eisenhardt 1990; Shepherd et al. 2015). Managers typically predict possible options by collecting and evaluating all relevant information and attempting to make decisions they perceive will best maximize their economic returns. However, managers often do not follow the procedures for rational choices (Mousavi and Gigerenzer 2014; Simon 1955). Rather, they rely on their heuristics, a simple decision-making process that utilizes only a fraction of the available information (Bingham et al. 2007; Bingham and Haleblian 2012). Research on managerial decision-making has also shown that making consistently optimal decisions is difficult due to bounded rationality (Simon 1991), cognitive biases (Thaler 1993), or perceptions deviating from economic optimality (Kahneman 2003). To mitigate these biases and errors, researchers propose to set goals and aspirations to guide decision-making and to use backward- and forward-looking decision models (Chen 2008; Cyert and March 1963; Gavetti and Levinthal 2000).

Information technology (IT) literature provides yet another set of solutions and argues that the adoption and utilization of new technologies compensate for these shortcomings. Information theory (Blackwell 1953) and the information-processing view of the organization (Galbraith 1974) propose that the more accurate and precise the information used in decision-making, the more the data used in decision-making leads to higher firm performance. This is primarily because IT

complements organizational practices, which leads to higher productivity (e.g., Bapna et al. 2013; Bresnahan et al. 2002; Brynjolfsson and Hitt 2000); IT also improves a firm's ability to collect, analyze, and process information for internal operational decisions. The positive relationship between the quality of information and decision-making has been supported by a number of empirical studies (e.g., Ayres 2007; Brynjolfsson et al. 2011; Davenport and Harris 2017; Loveman 2003).

As data availability has grown, researchers have extended these arguments to data-driven decision-making. The availability of data about consumers, suppliers, competitors, and partners and the utilization of large-scale analytics have supported managerial decision-making (Brynjolfsson et al. 2011; Wu et al. 2019). For example, Brynjolfsson et al. (2011) find that the adoption of data-driven decision-making practices increases financial returns. Saunders and Tambe (2013) find that firms with data-driven decision-making at an executive level have higher productivity and market valuations. Data analytics also support decision-making for R&D search and incremental process improvements (Wu et al. 2020). Overall, the adoption of new IT plays an important role in decision-making at both organizational and individual levels.

Researchers have recently extended this discussion to the adoption and utilization of AI. The advance in AI stemming from the development of machine learning and deep learning algorithms contributes to the avoidance of mistakes and errors stemming from human judgments that were previously regarded as relatively accurate (Danziger et al. 2011). AI algorithms are fundamentally different from previous algorithmic technologies for several reasons (Agrawal et al. 2018; Smith 2019). First, AI improves performance through self-learning, making inferences on new data based on training. Due to the ability to reason, AI can conduct insightful tasks that need human-like "intuition." Second, AI obtains predictions and judgments with high accuracy through

almost unlimited training based on voluminous data. Third, AI algorithms are optimized to predict the best decisions. These distinct characteristics enable AI to outperform humans not only in manual work and routine cognitive tasks but also in creative tasks in some domains (He et al. 2015; Mnih et al. 2015). Researchers find that AI performs well in high-level cognitive tasks such as making a legal decision in the court (Kleinberg et al. 2017), discovering protein structure in biology (Senior et al. 2020), and playing strategic games (Schrittwieser et al. 2020), among others.

Organizations therefore have begun to use AI algorithms to improve allocating valuable resources, designating work schedules, and analyzing employee performance (De Cremer, 2020). For example, AI algorithms are helpful in discovering new drugs—in particular, at the early stage when the tasks are heavily dependent on automatic data processing and reasoning (Lou and Wu 2020). Medical workers in hospitals also use AI suggestions in the chart coding area and thereby improve productivity (Wang et al. 2019).

In particular, considering the assumption of bounded rationality—that decision-makers tend to balance the accuracy of their decisions with the cost, such as the cognitive effort and time required to reach their decisions (Kahneman 2003)—AI can contribute to lower cost but can also provide better choices. AI helps the process of human decision-making by evaluating a broader scope of options at a lower cost, by providing a better evaluation of the options, and by providing the optimal choice among the options available. Therefore, our broad prediction is that the introduction of AI to human professionals improves the quality of human decision-making—especially when AI performance is superior to that of humans.

### 2.2. Junior vs. Senior AI utilization

Studies have found that age is an important factor in adopting new technology (Weinberg 2004). Age could affect the adoption and utilization of AI through three major channels: risk-aversion,

search behavior, and absorptive capacity. First, junior (younger) professionals are more open to new technologies and less risk-averse than senior professionals are. Hambrick and Mason (1984) proposed that an executive's age affects strategic decisions; for example, younger managers pursue riskier strategies in terms of the usage and development of new technology (Tyler and Steensma 1998). Given that technologies based on AI are risky and that this brand-new technology may be unproven and lack a track record, junior professionals are more inclined to adopt AI.

Second, senior professionals, who have more experience and accumulated knowledge, are more dependent on these compared to junior professionals. In other words, senior professionals conduct exploitative searches and make *path-dependent* choices based on past records. On the other hand, junior professionals without much experience are more likely to explore and make *path-creating* decisions, including adopting and utilizing AI.

Third, younger professionals tend to have better absorptive capacity (Cohen and Levinthal 1990) for new technology. Empirical evidence indicates that younger workers tend to adopt and to be more qualified to employ information and communications technologies while older workers are less likely and less qualified to use these (de Koning and Gelderblom 2006; Meyer 2011; Morris and Venkatesh 2000; Schleife 2006). The young are better able to recognize the value of new technology, to assimilate it, and to apply it to their professional tasks.

All the above arguments suggest that the impact of AI on human decision-making varies between junior and senior professionals and that junior professionals benefit most from this new technology.

## 3. EMPIRICAL STRATEGY

*"In a short space of time, AlphaGo Zero has understood all of the Go knowledge that has been accumulated by humans over thousands of years of playing … it's actually chosen to go beyond*

*that and discovered something that the humans hadn't even discovered in this time period."*

– David Silver, Google's DeepMind Lab

## 3.1. Setting

### *A. The game of GO*

Go (or Baduk) is a two-player strategy board game that originated in China at least 3,000 years ago. The board consists of nineteen lines by nineteen lines. Players compete to obtain more of the board's territory by alternating the placement of stones at the intersection of the lines. The professional Go market is substantial—especially in China, Japan, South Korea, and Taiwan. Each country holds more than ten major professional tournaments, held throughout per year and sponsored by large corporations. For example, the Kisei tournament in Japan—held annually since 1977 and sponsored by the Yomiuri Shimbun newspaper—awards 4,500,000 yen (or $413,000) to the first-place winner, in addition to per-game compensations.[1]

Demis Hassabis, head of the Google DeepMind team, noted that "Go is the most complex and beautiful game ever devised by humans … the richest in terms of intellectual depth." Go has about $250^{150}$ possible moves, and the search space is often described as "a number greater than there are atoms in the universe" (Silver, 2016).[2] The seemingly unlimited number of possible moves in Go cannot be exactly identified by brute force calculation (which supercomputers have done with chess); in the past two decades, several Go software programs—such as Tygem, GnuGo, Pachi, and Crazy Stone—were released, but the performance of these programs was far inferior to

[1] Other examples of major competitions include the Nongshim Cup, the competition between Team China, Japan, and South Korea, which awards $450,000 to the winning team. The Ing Cup (also known as Go Olympics), is held every four years and awards $400,000 to the winning player. In 2020, Jinseo Shin (a twenty-one-year-old) from South Korea earned $920,754 in award money; Imaya Yuta (a thirty-year-old) from Japan earned $1,179,456.
[2] For comparison, chess has about $35^{80}$ possible moves. After the first two moves, chess has 400 possible next moves, while Go has 130,000 possible next moves (Muoio 2016).

that of professional Go players who use superlative "intuition" and evaluation skills in making a certain move (Knight, 2016).

***B. AI came to Go***

Even if the latest supercomputers cannot calculate all possible moves in Go, the recent advancement in deep learning algorithms in AI has made a remarkable improvement. Instead of evaluating all possible solutions, AI uses deep learning to reduce the potential moves to be considered and predicts sequential outcomes and winning probabilities. AlphaGo, the initial APG with these algorithms, was invented by Google DeepMind. After several quality tests, Google held a historic Go match in 2016 between AlphaGo and the human Go master, Sedol Lee. Prior to this match, Lee and other Go experts expected that Lee would sweep all five games. Yet the result was that AlphaGo beat Lee 4–1, "a feat previously thought to be at least a decade away" (Silver et al., 2016). This event has been described as one of the milestones in the history of AI (Press 2021).

The result shocked not only Go players but also the public who believed computers to be far interior at intuitive judgments made amid enormous complexity. The match suddenly and unexpectedly demonstrated that AI-powered Go software surpassed the best human Go player. The match completely changed how professional Go players learn and practice Go; since the release in 2017 of public APGs—such as Leela Zero, KataGo, and Handol—all professional players learn from APG (Somers 2018).

***C. How much better at Go is AI compared to humans?***

Go players are ranked and evaluated by the Elo rating system.[3] Figure 1 shows how Elo scores have evolved between Go players and APGs. Non-AI Go software—GnuGo, Pachi, and Crazy Stone—scored under 2,000. The best human players scored around 3,800. In contrast, the scores

[3] The Elo rating is calculated based on the relative capabilities of two players and their game outcome. The system has been widely used in other sports such as chess, football, basketball, and soccer.

of APGs based on deep reinforcement learning far exceed 4,000. Given this gap in the Elo ratings, even the top professional Go players have no chance of winning against APGs. Ke Jie, who ranked second in the 2020 World Go Ranking, admitted that "AlphaGo is more like the god of Go." Put differently, the moves by APG yield the highest probability of winning and even the best professional Go player could learn a lot from APGs.

*"Insert Figure 1 here"*

***D. How does Go resemble managerial decision-making?***

The decisions made in each move in Go share many aspects with managerial decision-making. Studies have emphasized the link between decision-making in games and real businesses. For example, Simon (1987) argues that the intuitive skills required for managers are similar to the intuitive skills of chess masters. Mintzberg (1973, 1990) finds that half of the activities of five top executives lasted less than nine minutes. US supervisors perform on average 583 activities per eight-hour shift and perform one activity every 48 seconds (Guest, 1956). These executives and supervisors rely on intuitions and routines when making decisions (Mintzberg, 1973). Likewise, professional Go players face a complex, competitive environment and are forced to make a series of important decisions within a strict time restriction; after using up the regular time, they must make each move within thirty seconds.

Executives and practitioners also confirm this point. John Koo, chairman of LS Future Center and a pan-LG group family member, noted that "Baduk is a battle that starts out from a small part of the board and later expands to the entire board. You need to make your move while seeing the bigger picture from the very beginning. Business management is the same *(Korea Herald*, 2014)." LG Economic Research Institute(2004) also published a report, *Learning Business Strategy from the Principles of Go*, highlighting that both executives and professional Go players

must make decisions ceaselessly under uncertainty.

### 3.2. Research design

We study how professional Go players have learned from APG since its first public release. Although AlphaGo was the first APG to beat the best professional Go player, in 2016 only a scientific article about its algorithm—not the program itself—was not available to players. Leela, released in 2017, was the first APG available to professional Go players. The developer, Gian-Carlo Pascutto, made it clear that he invented the program independent of the needs of Go players (and he had no interest in playing Go); he just wanted to learn how deep learning works. Yet, the impact of Leela, which provided a set of best possible moves with the winning probability of each alternative— was substantial for Go players. Particularly, Leela and Leela Zero are the world's most successful open source Go engines based on AI (Somers 2018). We describe how Leela and Leela Zero work in Appendix A.

We use the event-study method to estimate the impact of APG on the quality of moves by professional Go players. The event of interest is the introduction of Leela as the first APG, which far outperformed the best human player. We conduct the analyses either at the game level or at the player-game level. Our sample consists of major professional Go games held from January 2015 to December 2019.

We focus mainly on early moves—up to 30 moves for each game—because, like many other games, starting with a great opening in Go is critical to winning the game. Chang-ho Lee, a legendary Go player, pointed out the importance of the opening and likened it to a blueprint for architecture. The opening strategies are general roadmaps and ideas on how players lead the game. As such, the first thirty moves are where professional Go players accumulated most of their knowledge and wisdom. Several opening sequences (*joseki)* have been long established as standard

procedure and players deviate from these standards as the game proceeds. Hence, the early stage of the game is where APGs could potentially challenge the knowledge accumulated historically by professional players. We also analyze later stages and present the results.

### 3.3 Data

#### *A. Go games and players*

We collect data on professional Go games played from 2015 through 2019 from the Go4Go database, which has been widely used in studies on Go (e.g., Chao et al. 2018; Ramon and Struyf 2003; Wu et al. 2018). The data contains detailed information on the game, its players, Komi (the number of bonus points given to the second mover), the sequence of all moves, and the game outcome. From Go Ratings we gathered additional data on the ages, nationalities (e.g., China, Japan, South Korea, Taiwan, and others), genders, and annual rankings of professional players. We use the negative value of ranking and divide it by 1,000 to ease the interpretation of the result; the higher the value, the better the player. To control for the difference in players' capabilities for each game, we create a variable, *Rank difference*, as the difference between the raw rankings of two players; we divide this difference by 1,000 such that a positive value indicates that the focal player's ranking is lower than the opponent's ranking.

#### *B. Measuring the quality of moves*

To evaluate the quality of moves by professional Go players, we use Leela Zero as a benchmark. Leela Zero is one of the highest performing APGs and is widely used by professional players and the public. For example, the Korea Baduk (Go) Association and the South Korean National Go Team use Leela Zero for learning and training. Because Leela Zero provides the probability of winning at any particular point of the game and the changes in the probability of winning for any move, we can compare the changes in winning probability between a move made by a professional

player and Leela Zero's suggested move, which increases the winning probability more than any other alternative move. For instance, if a player makes the same move as the one suggested by Leela Zero, the difference in winning probability between the two is zero—that is, the move has the highest winning probability among all alternatives. Thus, a smaller difference between a player's moves and those of Leela Zero indicates higher-quality moves by the professional Go player.

Our main dependent variable is *Diff from* $APG_{ig}$, representing the average difference in winning probability between a player's move and Leela Zero's suggestion for the first thirty moves in a game between two professional players. For each game, we calculate the value for the first mover (black stone holder) and the second mover (white stone holder), respectively. We used Leela Zero (May 23, 2020 version) along with the GoReviewPartner program to analyze all 25,033 games played from 2015 through 2019. Using two to eight Nvidia Titan-X GPUs running in parallel, the game analysis took about three months. In Appendix A, we describe how Leela Zero provides in-depth analysis to a game between two professional players. The implementation and calculation details are provided in Appendix B.

### *C. Summary statistics*

Table 1 provides descriptive statistics on the key variables at the player-game and player levels. Table 1(a) lists two observations for each game: one for the first mover (black stone holder) and another for the second mover (white stone holder). After games that lack information on players' ages or ranks are deleted, our final sample has 46,454 observations. The mean of our main dependent variable, *Diff from APG*, is 2.01. This means that the winning probability of professional player's first 30 moves averages 2.01 percentage points less than that of APG's suggestion. This is a huge difference because the winning probability of *each move* of the professional is 2

percentage points lower than that of the AI's best solution. The average (raw) rank of the players is 280th before transformation. The average rank difference is, by definition, zero (the positive and negative differences of the two players cancel out).

Table 1(b) shows the descriptive statistics at the player level. We identified 1,242 professional players from Go matches from 2015 through 2019. Due to the missing information on the ages and ranks of some players, our final sample contains 1,088 players. The average age of players is 32.73, and the median age is 27.32.

*"Insert Table 1 here"*

## 4. RESULTS

### 4.1. Does APG improve the quality of moves by professional players?

#### *A. Model-free evidence*

We first graphically present the average difference in winning probability between APG and professional Go players weekly from 2015 through 2019. The vertical line represents the times when the first APG, Leela, was released in February 2017. This model-free evidence shows that the difference in winning probability between APG and professional Go players was large and stable over time before the first release of APG. The gap decreases immediately after the public release of Leela in February 2017.

*"Insert Figure 2 here"*

#### *B. Event-study analysis*

We then use a formal OLS regression model to estimate the difference in the quality of moves between APG and professional players. The baseline event-study regression specification at the player-game level is:

$$Y_{ig} = \alpha + \beta_1 \cdot Post_g + \gamma_i + \delta_i + \epsilon_{ig}$$

where indices $i$ and $g$ represent player and game respectively. $\gamma_i$ represents focal player fixed effects while $\delta_i$ represents fixed effects for the opponent player. $Y_{ig}$ is the average difference in winning probability between the first thirty moves of the professional player and APG's best suggestion. $Post_g$ is equal to 1 if a Go game is played after the first public introduction of APG in February 2017 and 0 otherwise. Standard errors are clustered at the focal-player level to address a concern that the error terms are correlated across the players. $\beta_1$ captures how APG improved the moves by professional players.

The results are shown in Table 2. Column 1 shows that the coefficient of $Post_g$ is negative and significant ($\beta = -0.756, p < 0.01$), indicating that the difference in the winning probability between professional players and APG decreased by 0.756 percentage points (or about 30.5 percent) on average after APG's public release.

*"Insert Table 2 here"*

Yet, it is possible that professional players' performance had been improving over time and is driving the results, although Figure 2 does not indicate evidence of this. To control for this trend, we add a *Trend* variable (i.e., the number of quarters passed since the first quarter in our sample) and an interaction term (*Trend* × *Post*). The results are shown in column 2. We find a small yet negative trend ($\beta = -0.007, p < 0.05$), suggesting that the performance of professional players had been improving slowly over time. Importantly, the coefficient of the interaction term ($\beta = -0.116, p < 0.01$) shows that there were much larger (seventeen times greater) improvements following the public release of APG, even when the pre-APG performance trends are taken into account.

***C. Difference-in-differences estimation***

One may still worry that there was a general improvement in the performance of professional Go players around 2017. To address this concern, we further estimate the difference-in-differences model using country-level variations in adopting and utilizing APG. Among the three major countries holding the largest professional Go leagues—(mainland) China, Japan, and South Korea— Japan showed little awareness of or interest in APGs. For example, two historical matches between the countries' best players and APG were held in South Korea (AlphaGo vs. Sedol Lee, February 2016) and China (AlphaGo vs. Ke Jie, May 2017) but none was held in Japan. This reflects the countries' preexisting interest levels in APG, which further affected their post-event interests in it. Indeed, a search of Google Trends reveals that, from 2016 through 2017, the term *AlphaGo* was searched most by two countries: China and South Korea. In contrast, Japan was ranked seventh with an interest score of 5 (compared to China's score of 100 as a reference point and South Korea's score of 97). This motivates us to use the games by Japanese players as a counterfactual in the difference-in-differences estimation. We estimate the following:

$$Y_{ig} = \alpha + \beta_1 \cdot Post \cdot Treat_i + \beta_2 \cdot Treat_i + X_g + \gamma_i + \delta_i + \theta_g + \epsilon_{ig}$$

The treatment group consists of (mainland) China and Korea. The control group, Japan, takes into account any non-country-specific changes in performance of professional Go players. Although Japan is to some extent affected by APGs, this will bias our estimates toward zero (i.e., against our findings).

We present the results from difference-in-differences estimation in Table 3. Figure 4 illustrates the event study approach, with distributed leads and lags. In the main analysis, the difference decreased by 0.756 percentage points (Table 1, column 1). In the difference-in-differences model, as expected, the magnitude is smaller: –0.230 to –0.315 percentage points (Table 3, columns 2–4), statistically significant at the 0.01 level. Although to a lesser extent, this may be because our control group (Japan) also benefitted from the release of APG. That is, the

difference-in-differences estimate provides a lower bound estimate. From this stringent model, however, we once again confirm that AI is indeed what improves the quality of moves by professional players.

*"Insert Table 3 here"*

### 4.2. The differential effects of APG by age

As mentioned in Section 2.2, age is an important factor that could affect the adoption and utilization of new technology. We thus go further and test the differential effects of APG by player age. We estimate the following regression model at the player-game level:

$$Y_{ig} = \alpha + \beta_1 \cdot Post \cdot Young_i + \beta_2 \cdot Young_i + X_{ig} + \gamma_i + \delta_i + \theta_g + \epsilon_{ig}$$

where $\gamma_i$, $\delta_i$, and $\theta_g$ represent player, opponent player, and quarter-fixed effects, respectively for game $g$. $X_{ig}$ includes player or game-level control variables such as *Komi*, *White*, *Rank*, and *Rank* differences between players. *Young* is an indicator variable equal to 1 if the age of players is lower than the median age of all players (i.e., under twenty-eight years) at the date of APG's public release and 0 otherwise.

Table 4 shows the results. Column 1 includes only *Young* and control variables with quarter-time fixed effects. Column 2 then adds the interaction term, $Post \cdot Young$. The coefficient of the interaction term ($\beta = -0.273$, $p < 0.01$) is negative and significant; the quality improvement of moves by APG is 0.273 percentage points (or 11 percent) greater for junior players than for senior players.

To check whether our results hold if we consider the players' inborn characteristics, column 3 adds the player-fixed effect; column 4 adds the opponent player–fixed effects. Even after including the fixed effects for the players, we consistently find that the effect of AI is more

prominent for the junior group, which benefited 0.209–0.273 percentage points (8.5 percent–11 percent) more than the senior group. Additional robustness checks are provided in Section 4.3.

*"Insert Table 4 here"*

### 4.3. Robustness checks

We check the robustness of the results in five ways: 1) an estimation with distributed leads and lags, 2) a sensitivity test by age conditions, 3) a placebo test using the pseudo treatment assignment, 4) an analysis using monthly data, and 5) the different number of moves for an opening strategy (up to 40, 50, or 60 moves).

***A. Event study with distributed leads and lags***

To check the pre-APG trend and the time-varying effects of the APG, we include the distributed leads and lags in our regression and estimate the following:

$$Y_{ig} = \alpha + \Sigma_z \beta_t \cdot Z \cdot Young_i + X_{ig} + \gamma_i + \delta_i + \theta_g + \epsilon_{igt}$$

where $\gamma_i$, $\delta_i$, and $\theta_g$ represent player, opponent player, and time-quarter-fixed effects, respectively. $z$ represents the time leads and lags—that is, the number of quarters before or after the public release of APG.

*"Insert Figure 3 here"*

Figure 4 illustrates the results. We do not find any pre-APG trend; the estimates for pre-APG quarters are close to zero and not statistically significant; we cannot reject the null hypothesis that the estimate is equal to zero. For quarters after the public release of APG, the estimates are large and statistically significant. Table 5, columns 1–4, shows the detailed regression results.

*"Insert Table 5 here"*

***B. Sensitivity test for age groups***

We test whether the results are sensitive to our operationalization by age group. We conduct two

robustness checks using different age classifications and adding more granular age categories. First, we use the average age (instead of median age) as the cutoff for the junior versus senior group. The cutoff increases from twenty-eight years to thirty-three years. The results, provided in Appendix C, Table C.1, are robust to this alternative classification ($\beta = -0.289, p < 0.01$ in column 4).

In addition, we conduct the sensitivity test with more age groups. We investigate the same model with three age groups: those younger than twenty ("Young"), those in their twenties ("Middle"), and those thirty or older ("Old"). The results are provided in Appendix C, Table C.2. The estimates for $Young \cdot Post$ ($\beta = -0.302, p < 0.01$) and $Middle \cdot Post$ ($\beta = -0.198, p < 0.01$) are large and statistically significant. What is more, the effect is largest for Young players and decreases monotonically for Middle and then for Old players.

***C. Placebo test for age groups***

To check whether we capture the spurious variations when testing the age effects, we conduct a placebo test. We randomly reassign the age groups and estimate the models. If our suggested logics hold, we expect to find null effects—that is, we cannot reject the null hypothesis that the age effect is zero. As shown in Table C.3., Appendix C, the estimate for $Young \cdot Post$ is close to zero and *not* statistically significant with the randomly assigned age group.

***D. Alternative time-fixed effects and opening moves***

First, we estimate the model with month-fixed effects instead of quarter-fixed effects. Table C.4 of Appendix C shows that the results are consistent with this alternative time-fixed effects. Second, we re-estimated our models with different definitions of early moves: 40, 50, and 60 opening moves. The results, shown in Tables 5 of Appendix C, are robust to these alternative definitions, confirming that the operationalization of early moves does not drive the findings.

## 5. FURTHER ANALYSES

### 5.1. Quality improvement: Reducing errors and the critical mistake

In the main analysis, we look at the average difference between the moves by professional players and APG's best solution. In this section, we extend the analysis and consider two dimensions of move quality: *errors* and the *critical mistake*. We measure the error as the *number* of "bad" moves in the game where the winning probability drops by 10 percentage points or more compared to the previous move. The critical mistake is the *value* of the biggest drop in winning probability among all the moves in a game. Figure 5 shows the model-free trend of errors (in panel a) and the critical mistake (in panel b). Both errors and the critical mistake show a substantial decrease after the public release of APG.

*"Insert Figure 5 here"*

We then conduct regression analyses on errors and the critical mistake. Table 6, columns 1 and 3, show that the number of errors and the magnitude of critical mistakes decreased after APG release. Columns 2 and 4 show the results after controlling for the linear trend. The estimates for the interaction term ($Trend \cdot Post$) show that the negative trend is discontinuously accelerated after the introduction of APG. These results, again, confirm that AI improved professional players' moves by reducing both the frequency of errors and the magnitude of the critical mistake.

*"Insert Table 6 here"*

### 5.2. Do AI-driven improvements of moves lead to winning?

Junior players improve more than senior players after APG. We further investigate whether this move-level improvement—that is, a lower difference compared to AI's best solution, fewer errors, and smaller critical mistakes—leads to a higher probability of winning the game.

We conduct the three-step mediation analysis suggested by Baron and Kenny (1998). As a baseline model, we run a logit regression of being a junior player on winning after APG. We confirm from Table 7, column 1, that the estimate for $Post \cdot Young$ ($\beta = 0.120, p < 0.01$) is positive and statistically significant. The move-level improvements indeed lead to higher winning probability.

We then conduct the median analysis to test for the channels. The first step is to check whether $Post \cdot Young$ is statistically related to the proposed mediators: the quality difference (*Diff from APG*), errors (*number of bad moves*), and the magnitude of critical mistake (*Biggest probability dropped*). Table 7, columns 2–4, show that all three dependent variables are indeed negatively associated with the junior group after APG.

The second step is to check whether the mediators are negatively associated with the probability of wining, without the explanatory variable ($Post \cdot Young$). We confirm that this is the case from the results in Table 7, columns 5–7.

As the last step, we examine whether the magnitude of the estimated effect of the explanatory variable ($Post \cdot Young$) decreases with inclusion of the mediators. In Table 7, columns 8–11, the estimates for the explanatory variable ($Post \cdot Young$) decrease for all cases after adding the mediator variables, compared to those in the baseline model (column 1).

The mediation analysis confirms that junior players are more likely to win after APG through their improvements in three dimensions: quality difference compared to APG's best solution, the number of errors, and the magnitude of critical mistakes.

*"Insert Table 7 here"*

### 5.3. How does the AI effect vary throughout the game?

Although we focus on the early (first to thirtieth) moves in the main analyses, the role of AI is not

restricted to this particular phase. In this section, we extend the analysis to include later stages of the game. We add thirty moves incrementally to our baseline analysis—up to 180 moves—and compare the effects.

We first graphically present model-free results on the quality difference between professional players and APG's best solution (*Diff from APG*) in Figure 6. We find that the AI effect is most prominent at early opening moves (moves 1–30) and gradually decreases as we include later moves in the analysis.

*"Insert Figure 6 here"*

Formal analyses confirm these observations. Table 8 shows the results from six different regression specifications. The estimate for $Post \cdot Young$ gradually shrinks from –0.209 (for moves 1–30) to –0.053 (for moves 1–180). The event-study estimates with distributed leads and lags are graphically illustrated in Figure C.1, Appendix C. This clearly shows that junior players' improvement by APG is highest for the opening strategy and becomes weaker as moves from later stages of the game are included.

One explanation for this can be unexpectedness. At the early stage of a game, when only a few stones are placed, professional players can implement a prepared strategy learned from AI. As the game progresses into the mid-to-late stages, players face unexpected situations for which they are unprepared, where the learning from AI becomes less relevant; players may have to rely more on heuristics and routines to evaluate such unfamiliar environments. Another explanation is that, for later stages, more time is required to learn strategies from AI. For example, when professional players play 100 games with APG, they are likely to make very similar moves within the first thirty moves in all games (i.e., to learn 100 times). The game then develops into very different forms in later stages—that is, players learn that specific situation only once.

The results altogether suggest that AI's help with decision-making can vary depending on the uncertainty of the environment and the learning opportunity from—or exposure to—AI.

*"Insert Table 8 here"*

## 6. DISCUSSION AND CONCLUSIONS

> *Humans have evolved through thousands of years of actual combat exercises, but the AI-powered Go program has told us that humans are all wrong. Go players will combine computers to enter a whole new field and reach a whole new level.*
>
> —Ke Jie, Professional Go player (ranked second in World Go Ranking)

We study whether and how AI improves human decision-making. We exploit a unique setting, professional Go games, where AI algorithms clearly surpass the best human player and the knowledge humans have accumulated over thousands of years. We use the APG made public in 2017 to find the best move in each point of the game and measure the difference in the moves by professional players and AI's solution. This is a rare opportunity to observe every single decision made by human players and to evaluate the quality of these decisions. We find that AI algorithms improve the quality of moves by professional players (i.e., players are able to reduce the quality gap between their moves and those of AI) by decreasing the number of errors and the magnitude of the critical mistake. The effect is more prominent for the early stages of the game where players can implement the prepared strategy suggested by AI. Importantly, the junior group benefited more than the senior group.

The findings from professional Go players provide important implications for AI and human knowledge and skills. First, AI reveals that what humans believe to be the best may not actually be the best; AI could bring breakthroughs in human knowledge, heuristics, or routines that have been developed and improved for a long time. The application of AI, therefore, should have substantial effects for the practices and performance of individuals and organizations. As AI

outperforms the best human professionals in the job domain, it can pave a way for new paradigms. Relatedly, fast adoption of AI could increase the competitiveness of human professionals. In the professional Go context, AI has not replaced professional Go players; rather, it is used as a tool to improve the quality of human moves. Given the widespread fear that AI will replace human jobs, it is important to carefully study the areas where AI surpasses humans and to find ways to utilize AI to complement human jobs (or to help humans focus on more creative or value-added tasks). Second, not everyone may enjoy the benefits of AI. We show that the junior group benefited from APG. This suggests that the effects of AI may vary across groups, possibly by the groups' openness to new technology and their ability to utilize it. A recent study reveals that modern AI research and utilization is concentrated among elite universities and a small number of large corporate labs (Ahmed and Wahed, 2021). Thus it is important to carefully consider the differential access to AI and the potential inequality that could arise from that differential. Third, the impact of AI also depends on the complexity and uncertainty of a situation. We find that, in the case of Go, AI-based improvement is most prominent in the early stages of the game. This boundary condition of the AI effect is consistent with the example of new drug discovery by AI (Lou and Wu, 2021). Generally speaking, some tasks or areas (e.g., when there is high complexity) can benefit a great deal from AI, while human professionals could do better in other situations. A uniform application of AI would not yield the best outcome and could lead to inefficient allocation of AI-related resources. A careful consideration of where to adopt AI and to what extent is therefore required.

This paper contributes to several literature streams. First, this work contributes to our understanding of the impact of AI on human decision-making. As the influence of AI increases, researchers have focused on the role of AI in supporting human judgment (Choudhury et al. 2020, Kleinberg et al. 2017). We show that AI can help reevaluate the choices we believed were correct

due to heuristics and routines and can reduce humans' mistakes and errors. Second, we contribute to the literature on digital economy and the economics of AI. Considering that economists have focused on AI as a tool for automating prediction (Agrawal et al. 2019; Brynjolfsson et al. 2018) and as a GPTs (Cockburn et al. 2019; Goldfarb et al. 2019; Trajtenberg 2018), our findings contribute by showing that AI can provide new knowledge that humans do not have or have not yet recognized. For example, in Go there are classic opening strategies (called *joseki*) that humans have believed to be the best norms for hundreds of years. In the past, we could not evaluate the quality of these strategies. However, due to the decreasing cost of prediction, we are now able to show that APGs can evaluate the classic strategies and provide new opening strategies if *joseki* do not have the best probability of winning. Third, we contribute to the literature on micro-foundations of strategy in that decisions or actions at the individual level shape strategy and performance at the organizational level (Eisenhardt et al. 2010). This work shows that the strategic choices suggested by AI improve the competitive advantage of professionals and affect the likelihood of winning the game. Lastly, we contribute to the innovation literature on how the adoption of new technology has different impacts on human capital (Agarwal and Prasad 1999, Choudhury et al. 2020, Miric et al. 2020). We show that the utilization of AI varies among different age groups and that the application range of AI is limited at the early stage of the game.

Our study is not without limitations, which can provide further avenues for future research. Notably, decisions in professional Go games are made under clear rules of the game. Managers may also face a set of rules and restrictions, but they may have much more discretion in making decisions. Still, given that codification of routines in a firm is important (e.g., Bingham et al. 1997, Foss 2003), we believe that the application of AI will make up a large part of those routines. For example, the domains where AI outperforms humans have broadened to include different

organizations such as hospitals (Wang et al, 2019), law firms (Kahn, 2021) and sports teams (Katwala, 2021). We hope that future study extends the scope of this line of research on how AI improves human decision-making in different contexts.

## REFERENCES


Agarwal R, Prasad J (1999) Are Individual Differences Germane to the Acceptance of New Information Technologies? *Decis. Sci.* 30(2):361–391.

Agrawal A, Gans J, Goldfarb A (2018) *Prediction machines: the simple economics of artificial intelligence* (Harvard Business Press).

Ayres I (2007) *Super crunchers: Why thinking-by-numbers is the new way to be smart* (Bantam Books).

Bapna R, Langer N, Mehra A, Gopal R, Gupta A (2013) Human Capital Investments and Employee Performance: An Analysis of IT Services Industry. *Manage. Sci.* 59(3):641–658.

Bingham CB, Eisenhardt KM, Davis JP (1997) Opening the Black Box of Organizational Expertise: Understanding What Firms Learn from their Process Experience and How that learning Unfolds over Time. *Work. Pap. Univ. Maryl.*

Bingham CB, Eisenhardt KM, Furr NR (2007) What makes a process a capability? Heuristics, strategy, and effective capture of opportunities. *Strateg. Entrep. J.* 1(1–2):27–47.

Bingham CB, Haleblian JJ (2012) How firms learn heuristics: Uncovering missing components of organizational learning. *Strateg. Entrep. J.* 6(2):152–177.

Blackwell D (1953) Equivalent Comparisons of Experiments. *Ann. Math. Stat.* 24(2):265–272.

Bresnahan TF, Brynjolfsson E, Hitt LM (2002) Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence. *Q. J. Econ.* 117(1):339–376.

Brynjolfsson E, Hitt LM (2000) Beyond Computation: Information Technology. *J. Econ. Perspect.* 14(4):23–48.

Brynjolfsson E, Hitt LM, Kim HH (2011) Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance? *SSRN Electron. J.*

Chen WR (2008) Determinants of firms' backward- and forward-looking R&D search behavior. *Organ. Sci.* 19(4):609–622.

Choudhury P, Starr E, Agarwal R (2020) Machine learning and human capital complementarities: Experimental evidence on bias mitigation. *Strateg. Manag. J.* 41(8):1381–1411.

Cohen WM, Levinthal DA (1990) Absorptive Capacity: A New Perspective on Learning and Innovation. *Adm. Sci. Q.* 35(1):128.

Cyert RM, March JG (1963) *A behavioral theory of the firm*

Dane E, Pratt MG (2007) Exploring Intuition and its Role in Managerial Decision Making. *Acad. Manag. Rev.* 32(1):33–54.

Davenport T, Harris J (2017) *Competing on analytics: Updated, with a new introduction: The new science of winning* (Harvard Business Press).

Eisenhardt KM (1990) Speed and Strategic Choice: How Managers Accelerate Decision Making. *Calif. Manage. Rev.* 32(3):39–54.

Eisenhardt KM, Furr NR, Bingham CB (2010) CROSSROADS—Microfoundations of Performance: Balancing Efficiency and Flexibility in Dynamic Environments. *Organ. Sci.* 21(6):1263–1273.

Foss NJ (2003) Bounded rationality and tacit knowledge in the organizational capabilities approach: an assessment and a re-evaluation. *Ind. Corp. Chang.* 12(2):185–201.

Galbraith JR (1974) Organization Design: An Information Processing View. *Interfaces (Providence).* 4(3):28–36.

Gavetti G, Levinthal D (2000) Looking forward and looking backward: Cognitive and experiential search. *Adm. Sci. Q.* 45(1):113–137.

Kahneman D (2003) Maps of Bounded Rationality: Psychology for Behavioral Economics. *Am. Econ. Rev.* 93(5):1449–1475.

Kalberg S (1980) Max Weber's Types of Rationality: Cornerstones for the Analysis of Rationalization Processes in History. *Am. J. Sociol.* 85(5):1145–1179.

Kleinberg J, Lakkaraju H, Leskovec J, Ludwig J, Mullainathan S (2017) Human Decisions and Machine Predictions. *Q. J. Econ.* (January):237–293.

de Koning J, Gelderblom A (2006) ICT and older workers: No unwrinkled relationship. *Int. J. Manpow.* 27(5):467–490.

Korean Herald (2014) [SUPER RICH] Game of go makes great businesspeople. *The Korea Herald* (December 8).

Levinthal DA, March JG (1993) The myopia of learning. *Strateg. Manag. J.* 14(S2):95–112.

LGERI (LG Economic Research Institute) (2004) Learning Business Strategy from the Principles of Go. *LG Economic Research Institute (LGERI)*

Lindebaum D, Vesa M, den Hond F (2020) Insights From "The Machine Stops" to Better Understand Rational Assumptions in Algorithmic Decision Making and Its Implications for Organizations. *Acad. Manag. Rev.* 45(1):247–263.

Loveman G (2003) Diamonds in the Data Mine. *Harvard Buiness Rev.*

Meyer J (2011) Workforce age and technology adoption in small and medium-sized service firms. *Small Bus. Econ.* 37(3):305–324.

Mintzberg H (1987) *Crafting strategy: Harvard Business School Press Boston*

Mintzberg H (1994) Rethinking strategic planning part I: Pitfalls and fallacies. *Long Range Plann.* 27(3):12–21.

Miric M, Lu J, Teodoridis F (2020) Decision-Making Skills in an AI World: Lessons from Online Chess. *SSRN Electron. J.*:1–43.

Morris MG, Venkatesh V (2000) Age Differences in Technology Adoption Decisions: Implications for a Changing Work Force. *Pers. Psychol.* 53(2):375–403.

Mousavi S, Gigerenzer G (2014) Risk, uncertainty, and heuristics. *J. Bus. Res.* 67(8):1671–1678.

Press G (2021) 114 Milestones In The History Of Artificial Intelligence (AI). *Forbes* (May 19).

Saunders A, Tambe P (2013) A Measure of Firms' Information Practices Based on Textual Analysis of 10-K Filings. *Work. Pap.*

Schleife K (2006) Computer Use and Employment Status of Older Workers - An Analysis Based on Individual Data. *Labour* 20(2):325–348.

Schrittwieser J, Antonoglou I, Hubert T, Simonyan K, Sifre L, Schmitt S, Guez A, et al. (2020) Mastering Atari, Go, chess and shogi by planning with a learned model. *Nature* 588(7839):604–609.

Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, Green T, Qin C, et al. (2020) Improved protein structure prediction using potentials from deep learning. *Nature* 577(7792):706–710.

Shepherd DA, Williams TA, Patzelt H (2015) Thinking About Entrepreneurial Decision Making. *J. Manage.* 41(1):11–46.

Simon HA (1955) A Behavioral Model of Rational Choice. *Q. J. Econ.* 69(1):99.

Simon HA (1991) Bounded Rationality and Organizational Learning. *Organ. Sci.* 2(1):125–134.

Simon JGMHA (1958) *Organizations* (Wiley-Blackwell).

Smith BC (2019) *The promise of artificial intelligence: reckoning and judgment* (Mit Press).

Somers J (2018) How the Artificial-Intelligence Program AlphaZero Mastered Its Games. *The New Yorker* (December 28).

Thaler R (1993) *Advances in Behavioral Finance Russell Sage Foundation* (New York).

Tyler BB, Steensma HK (1998) The effects of executives' experiences and perceptions on their assessment of potential technological alliances. *Strateg. Manag. J.* 19(10):939–965.

Wang W, Gao G (Gordon), Agarwal R (2019) Friend or Foe? The Influence of Artificial Intelligence on Human Performance in Medical Chart Coding. *SSRN Electron. J.*

Weinberg B (2004) Experience and Technology Adoption. *IZA Discuss. Pap.* (1051).

Wu L, Hitt L, Lou B (2020) Data Analytics, Innovation, and Firm Productivity. *Manage. Sci.* 66(5):2017–2039.

**Figure 1:** Elo rating of AI-based Go Programs (APGs)

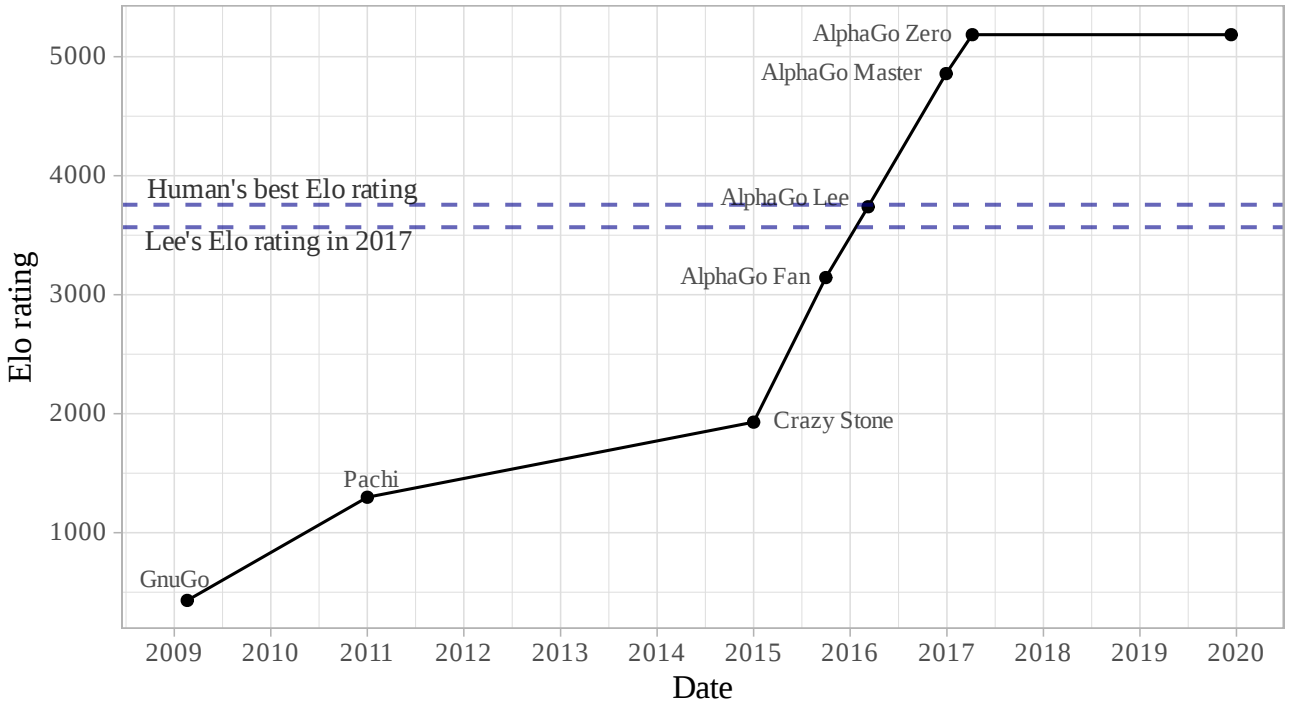


**Figure 2:** Average difference in winning probability between moves by players and APG's best solution: Model-free evidence

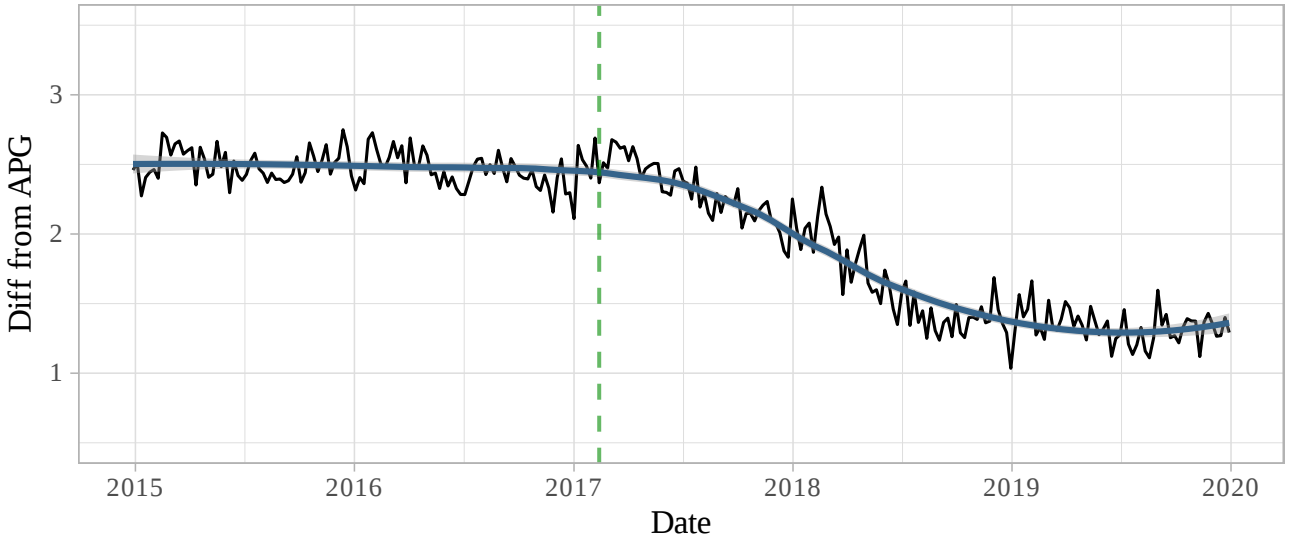



**Figure 3:** Difference-in-differences estimation: China and Korea (Treated) versus Japan (Control)

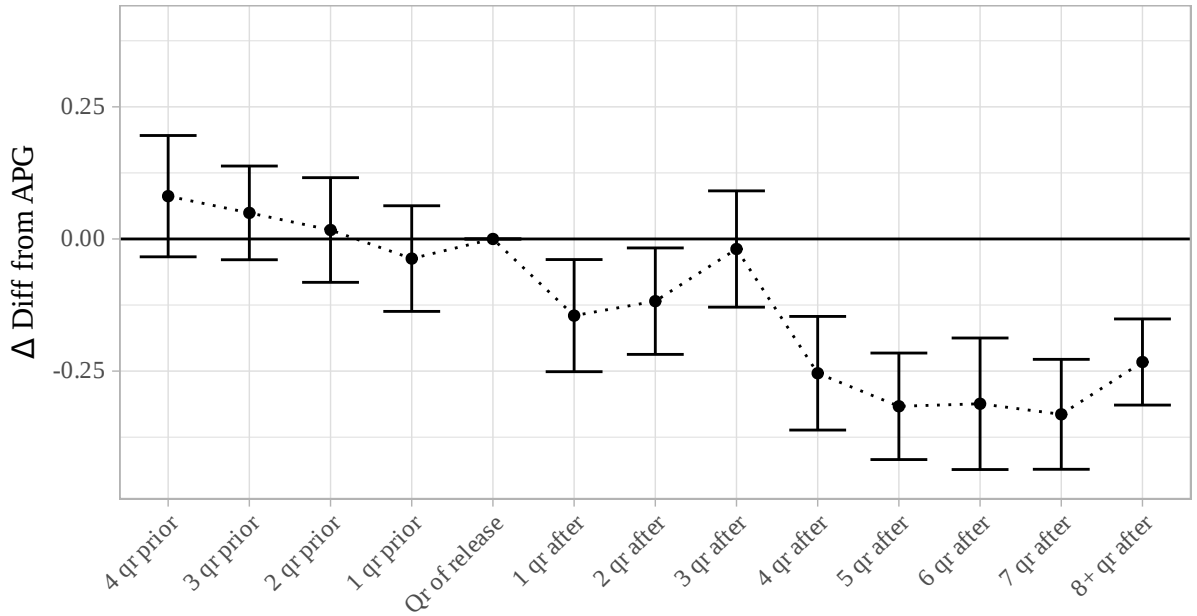
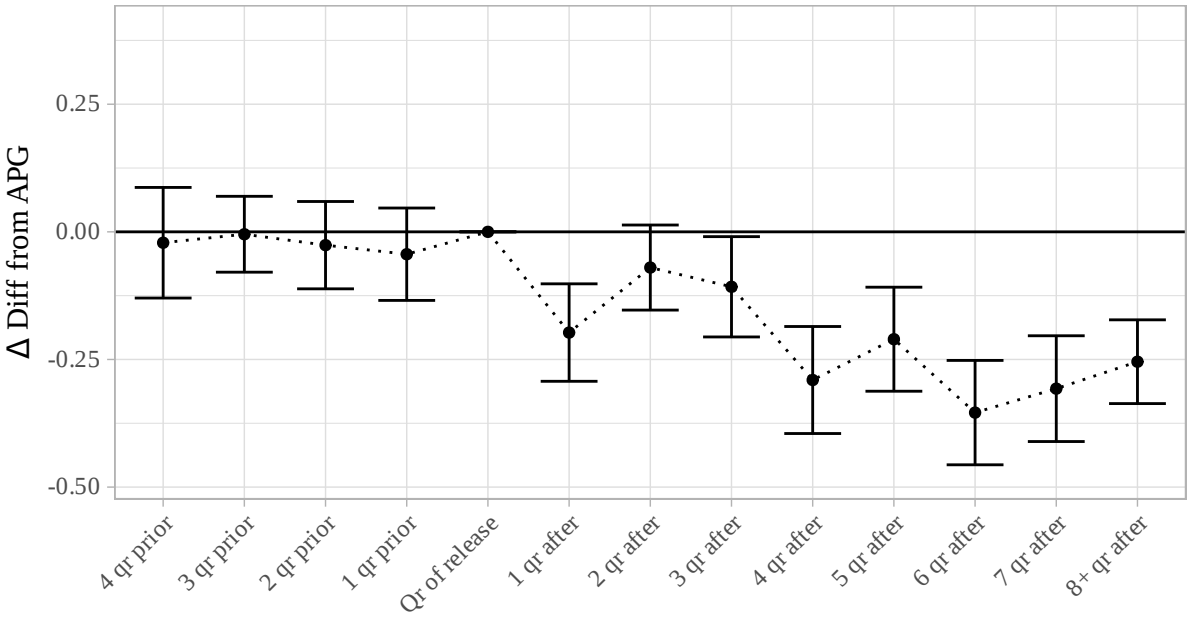
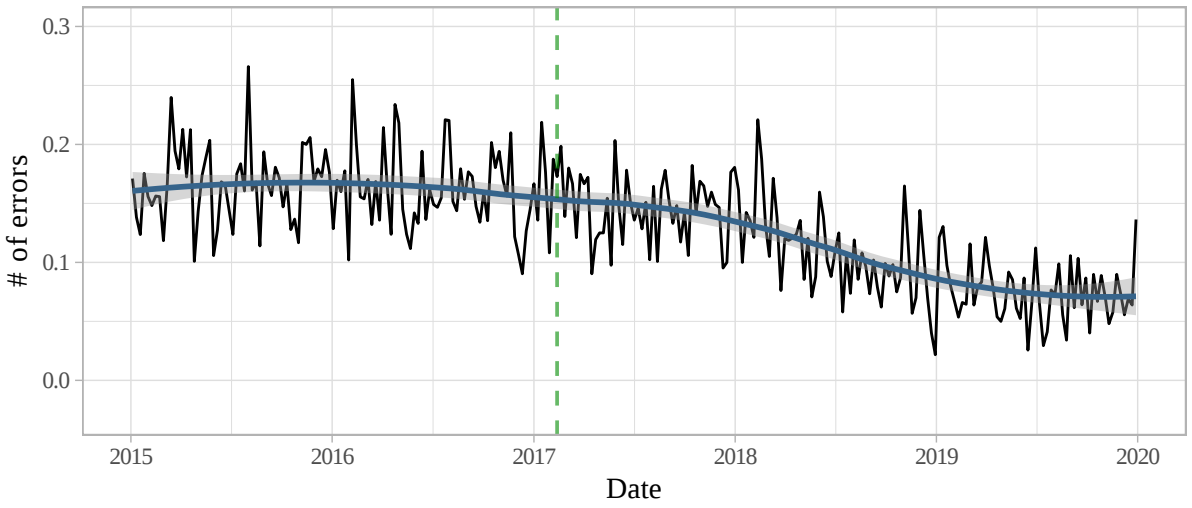


**Figure 4:** Differential effects by age:
Difference from APG for junior group compared to that of senior group

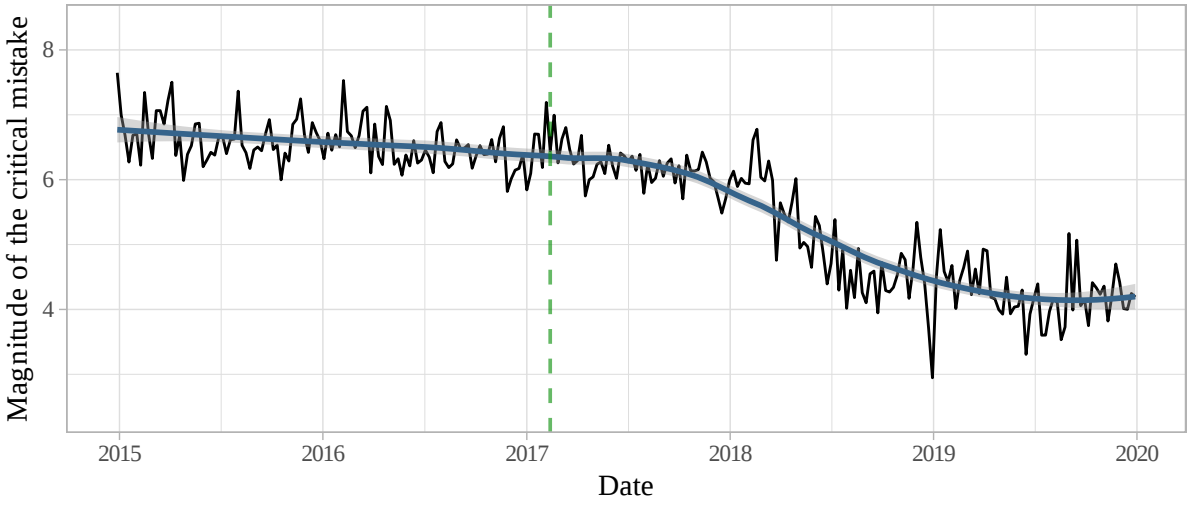



Time passage relative to the quarter of the release of APG

**Figure 5:** Errors and a critical mistake by professional Go players: Model-free evidence




**(a)** Number of errors




**(b)** Magnitude of the most critical mistake

**Figure 6:** Average difference in winning probability: Heterogeneity by the number of moves

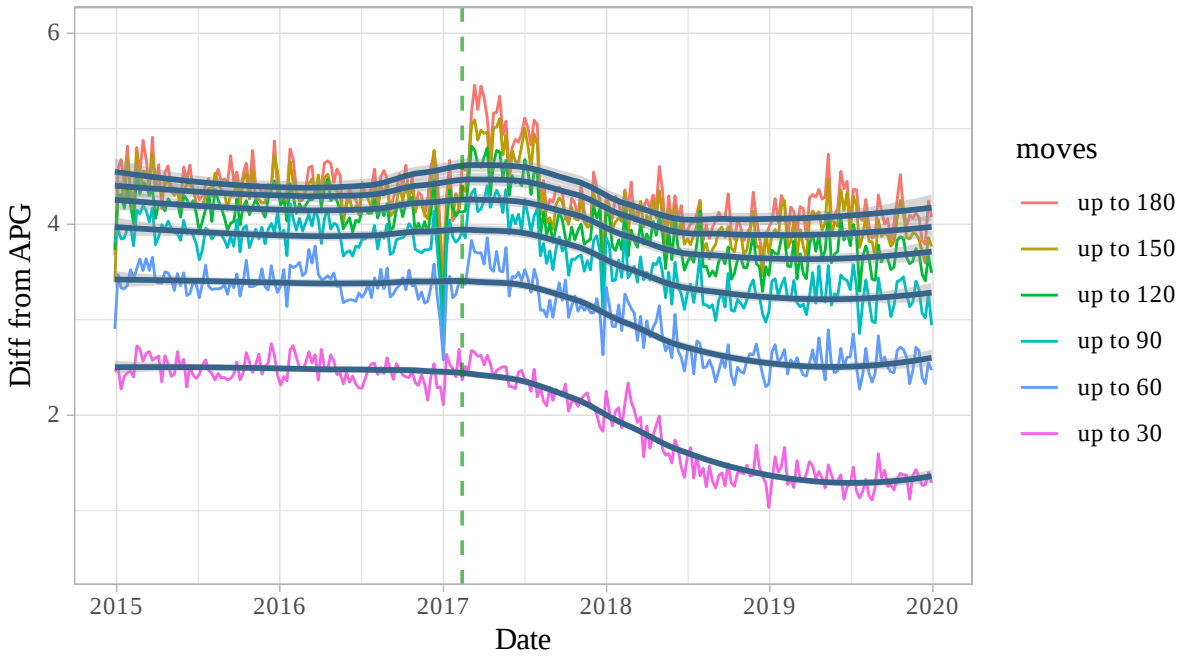


**Table 1:** Descriptive Statistics

| | N | Mean | Median | SD | P25 | P75 |
|---|---|---|---|---|---|---|
| Diff from APG | 50066 | 2.01 | 1.92 | 1.07 | 1.22 | 2.66 |
| Age | 49214 | 28.18 | 24.52 | 12.57 | 19.61 | 31.60 |
| Young | 49214 | 0.62 | 1.00 | 0.49 | 0.00 | 1.00 |
| Rank (-) | 48808 | -0.28 | -0.18 | 0.27 | -0.43 | -0.05 |
| Rank Diff (p-o) | 47712 | 0.00 | 0.00 | 0.22 | -0.09 | 0.09 |
| White | 50066 | 0.50 | 0.50 | 0.50 | 0.00 | 1.00 |
| 7.5 komi | 50066 | 0.38 | 0.00 | 0.49 | 0.00 | 1.00 |

Note: We use the negative value of ranking for the ease of interpretation and divide it by 1,000 (to scale the coefficient). *Rank Diff* rank differences as the relative difference between my ranking at the beginning of the year from the opponent. This value is also divided by 1,000 (to scale the coefficient)

**(a)** Player-game level

| | N | Mean | Median | SD | P25 | P75 |
|---|---|---|---|---|---|---|
| Diff from APG | 1242 | 2.20 | 2.18 | 0.67 | 1.79 | 2.52 |
| Age | 1149 | 32.73 | 27.32 | 16.15 | 20.25 | 42.77 |
| Young | 1149 | 0.49 | 0.00 | 0.50 | 0.00 | 1.00 |
| Rank (-) | 1088 | -0.54 | -0.54 | 0.31 | -0.81 | -0.27 |
| Rank Diff (p-o) | 1080 | 0.15 | 0.11 | 0.19 | 0.00 | 0.27 |

**(b)** Player level

**Table 2:** Difference in winning probability between moves by professional players and APG: Before and after the release of APG

| Dependent Variable: | Diff from APG | |
|---|---|---|
| Model: | (1) | (2) |
| *Variables* | | |
| Post | -0.756*** | 1.932*** |
| | (0.017) | (0.065) |
| Trend | | -0.007** |
| | | (0.003) |
| Trend × Post | | -0.116*** |
| | | (0.004) |
| *Fixed-effects* | | |
| Player | Yes | Yes |
| Opponent Player | Yes | Yes |
| *Fit statistics* | | |
| Observations | 50,066 | 50,066 |
| $R^2$ | 0.264 | 0.330 |
| Within $R^2$ | 0.116 | 0.195 |

*Clustered standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

**Table 3:** Difference-in-differences estimation:
China and Korea (Treated) versus Japan (Control)

| Dependent Variable: | Diff from APG | | | | | |
|---|---|---|---|---|---|---|
| Model: | (1) | (2) | (3) | (4) | (5) | (6) |
| *Variables* | | | | | | |
| Treat | -0.110*** | 0.069*** | | | | |
| | (0.024) | (0.026) | | | | |
| Rank (-) | -0.809*** | -0.810*** | -2.222*** | -3.146*** | -2.176*** | -3.085*** |
| | (0.040) | (0.040) | (0.221) | (0.275) | (0.221) | (0.278) |
| Rank Diff (p-o) | -0.133*** | -0.131*** | -0.059** | -1.115*** | -0.058** | -1.086*** |
| | (0.032) | (0.032) | (0.027) | (0.178) | (0.027) | (0.180) |
| White | 0.129*** | 0.129*** | 0.128*** | 0.126*** | 0.128*** | 0.126*** |
| | (0.010) | (0.010) | (0.010) | (0.010) | (0.010) | (0.010) |
| 7.5 komi | -0.011 | -0.010 | -0.020 | -0.030 | -0.021 | -0.032* |
| | (0.018) | (0.017) | (0.017) | (0.020) | (0.016) | (0.019) |
| Treat × Post | | -0.315*** | -0.266*** | -0.230*** | | |
| | | (0.031) | (0.031) | (0.031) | | |
| 4 qr prior | | | | | 0.063 | 0.081 |
| | | | | | (0.057) | (0.059) |
| 3 qr prior | | | | | 0.029 | 0.049 |
| | | | | | (0.045) | (0.045) |
| 2 qr prior | | | | | 0.006 | 0.017 |
| | | | | | (0.050) | (0.051) |
| 1 qr prior | | | | | -0.046 | -0.037 |
| | | | | | (0.052) | (0.051) |
| 1 qr after | | | | | -0.172*** | -0.145*** |
| | | | | | (0.053) | (0.054) |
| 2 qr after | | | | | -0.159*** | -0.118** |
| | | | | | (0.051) | (0.051) |
| 3 qr after | | | | | -0.048 | -0.019 |
| | | | | | (0.055) | (0.056) |
| 4 qr after | | | | | -0.291*** | -0.254*** |
| | | | | | (0.055) | (0.055) |
| 5 qr after | | | | | -0.361*** | -0.317*** |
| | | | | | (0.052) | (0.051) |
| 6 qr after | | | | | -0.354*** | -0.312*** |
| | | | | | (0.063) | (0.063) |
| 7 qr after | | | | | -0.387*** | -0.332*** |
| | | | | | (0.052) | (0.053) |
| 8 qr after | | | | | -0.278*** | -0.233*** |
| | | | | | (0.040) | (0.042) |
| *Fixed-effects* | | | | | | |
| Quarter | Yes | Yes | Yes | Yes | Yes | Yes |
| Player | | | Yes | Yes | Yes | Yes |
| Opponent Player | | | | Yes | | Yes |
| *Fit statistics* | | | | | | |
| Observations | 42,783 | 42,783 | 42,783 | 42,783 | 42,783 | 42,783 |
| $R^2$ | 0.277 | 0.281 | 0.327 | 0.352 | 0.328 | 0.353 |
| Within $R^2$ | 0.061 | 0.067 | 0.014 | 0.014 | 0.016 | 0.015 |

*Clustered standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

**Table 4:** Differential effects by age:
Difference from APG for junior group compared to that of senior groups

| Dependent Variable: | Diff from APG | | | |
|---|---|---|---|---|
| Model: | (1) | (2) | (3) | (4) |
| *Variables* | | | | |
| Young | -0.103*** | 0.049** | | |
| | (0.020) | (0.021) | | |
| Rank (-) | -0.817*** | -0.799*** | -1.626*** | -2.501*** |
| | (0.036) | (0.037) | (0.240) | (0.292) |
| Rank Diff (p-o) | -0.137*** | -0.128*** | -0.061** | -1.023*** |
| | (0.027) | (0.027) | (0.024) | (0.162) |
| White | 0.134*** | 0.133*** | 0.131*** | 0.131*** |
| | (0.010) | (0.010) | (0.009) | (0.010) |
| 7.5 komi | -0.025 | -0.023 | -0.024 | -0.041** |
| | (0.017) | (0.017) | (0.016) | (0.019) |
| Young × Post | | -0.273*** | -0.227*** | -0.209*** |
| | | (0.028) | (0.031) | (0.031) |
| *Fixed-effects* | | | | |
| Quarter | Yes | Yes | Yes | Yes |
| Player | | | Yes | Yes |
| Opponent Player | | | | Yes |
| *Fit statistics* | | | | |
| Observations | 46,454 | 46,454 | 46,454 | 46,454 |
| $R^2$ | 0.276 | 0.279 | 0.325 | 0.349 |
| Within $R^2$ | 0.064 | 0.069 | 0.013 | 0.014 |

*Clustered standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

**Table 5:** Average difference in winning probability between player's moves and APG's best solution: Event study estimates

| Dependent Variable: | Diff from APG | | | |
|---|---|---|---|---|
| Model: | (1) | (2) | (3) | (4) |
| *Variables* | | | | |
| Young | -0.103*** | 0.067*** | | |
| | (0.020) | (0.024) | | |
| Rank (-) | -0.817*** | -0.796*** | -1.429*** | -2.292*** |
| | (0.036) | (0.037) | (0.251) | (0.308) |
| Rank Diff (p-o) | -0.137*** | -0.127*** | -0.059** | -0.974*** |
| | (0.027) | (0.027) | (0.024) | (0.163) |
| White | 0.134*** | 0.133*** | 0.131*** | 0.131*** |
| | (0.010) | (0.010) | (0.009) | (0.010) |
| 7.5 komi | -0.025 | -0.024 | -0.025 | -0.043** |
| | (0.017) | (0.017) | (0.016) | (0.019) |
| 4 qr prior | | -0.038 | -0.041 | -0.021 |
| | | (0.055) | (0.054) | (0.055) |
| 3 qr prior | | -0.036 | -0.014 | -0.005 |
| | | (0.036) | (0.037) | (0.038) |
| 2 qr prior | | -0.017 | -0.030 | -0.026 |
| | | (0.043) | (0.043) | (0.044) |
| 1 qr prior | | -0.068 | -0.052 | -0.044 |
| | | (0.046) | (0.046) | (0.046) |
| 1 qr after | | -0.235*** | -0.209*** | -0.197*** |
| | | (0.049) | (0.048) | (0.049) |
| 2 qr after | | -0.109** | -0.088** | -0.070 |
| | | (0.045) | (0.043) | (0.043) |
| 3 qr after | | -0.132*** | -0.109** | -0.108** |
| | | (0.049) | (0.049) | (0.050) |
| 4 qr after | | -0.346*** | -0.303*** | -0.290*** |
| | | (0.054) | (0.054) | (0.053) |
| 5 qr after | | -0.299*** | -0.246*** | -0.210*** |
| | | (0.052) | (0.053) | (0.052) |
| 6 qr after | | -0.391*** | -0.376*** | -0.354*** |
| | | (0.050) | (0.052) | (0.052) |
| 7 qr after | | -0.384*** | -0.335*** | -0.307*** |
| | | (0.050) | (0.052) | (0.053) |
| 8+ qr after | | -0.333*** | -0.284*** | -0.254*** |
| | | (0.036) | (0.042) | (0.042) |
| *Fixed-effects* | | | | |
| Quarter | Yes | Yes | Yes | Yes |
| Player | | | Yes | Yes |
| Opponent Player | | | | Yes |
| *Fit statistics* | | | | |
| Observations | 46,454 | 46,454 | 46,454 | 46,454 |
| $R^2$ | 0.276 | 0.281 | 0.325 | 0.350 |
| Within $R^2$ | 0.064 | 0.071 | 0.015 | 0.015 |

*Clustered standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

**Table 6:** Errors and a critical mistake by professional Go players

| Dependent Variable: | Number of errors | | Magnitude of the critical mistake | |
|---|---|---|---|---|
| Model: | (1) | (2) | (3) | (4) |
| *Variables* | | | | |
| Post | -0.0516*** | 0.1552*** | -1.431*** | 4.124*** |
| | (0.0041) | (0.0245) | (0.0530) | (0.2552) |
| Trend | | -0.0004 | | -0.0280** |
| | | (0.0013) | | (0.0121) |
| Trend × Post | | -0.0091*** | | -0.2328*** |
| | | (0.0015) | | (0.0152) |
| *Fixed-effects* | | | | |
| Player | Yes | Yes | Yes | Yes |
| Opponent Player | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | |
| Observations | 50,066 | 50,066 | 50,065 | 50,065 |
| $R^2$ | 0.07666 | 0.08062 | 0.12348 | 0.14537 |
| Within $R^2$ | 0.00398 | 0.00825 | 0.02786 | 0.05215 |

*Clustered standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

**Table 7:** Mediation analysis on game winning: Difference in move quality, errors, and critical mistake

| Dependent Variables: | Win | Diff from APG | Number of errors | Magnitude of the critical mistake | | | | Win | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Model: | (1) Logit | (2) OLS | (3) OLS | (4) OLS | (5) Logit | (6) Logit | (7) Logit | (8) Logit | (9) Logit | (10) Logit | (11) Logit |
| *Variables* | | | | | | | | | | | |
| Rank (-) | -0.461 | -2.501*** | -0.369*** | -7.452*** | -0.548 | -0.220 | -0.427 | -0.790 | -0.546 | -0.712 | -0.838 |
| | (0.573) | (0.292) | (0.101) | (1.155) | (0.532) | (0.537) | (0.535) | (0.567) | (0.572) | (0.570) | (0.568) |
| Rank Diff (p-o) | -10.178*** | -1.023*** | -0.021 | -2.209*** | -10.333*** | -10.174*** | -10.276*** | -10.334*** | -10.195*** | -10.286*** | -10.355*** |
| | (0.490) | (0.162) | (0.065) | (0.683) | (0.483) | (0.484) | (0.486) | (0.490) | (0.491) | (0.494) | (0.493) |
| White | 0.125*** | 0.131*** | 0.047*** | 0.622*** | 0.146*** | 0.138*** | 0.148*** | 0.143*** | 0.135*** | 0.145*** | 0.149*** |
| | (0.025) | (0.010) | (0.004) | (0.037) | (0.025) | (0.025) | (0.025) | (0.025) | (0.025) | (0.025) | (0.025) |
| 7.5 komi | -0.001 | -0.041** | -0.011 | -0.102 | 0.001 | 0.004 | 0.004 | -0.005 | -0.002 | -0.002 | -0.004 |
| | (0.053) | (0.019) | (0.008) | (0.080) | (0.053) | (0.053) | (0.053) | (0.053) | (0.053) | (0.053) | (0.053) |
| Post × Young | 0.120*** | -0.209*** | -0.025*** | -0.508*** | | | | 0.094** | 0.115*** | 0.106** | 0.094** |
| | (0.044) | (0.031) | (0.008) | (0.101) | | | | (0.045) | (0.045) | (0.045) | (0.045) |
| Diff from APG | | | | | -0.132*** | | | -0.130*** | | | -0.084*** |
| | | | | | (0.012) | | | (0.012) | | | (0.015) |
| # of errors | | | | | | -0.206*** | | | -0.211*** | | 0.025 |
| | | | | | | (0.032) | | | (0.032) | | (0.045) |
| Magnitude of the critical mistake | | | | | | | -0.031*** | | | -0.031*** | -0.021*** |
| | | | | | | | (0.003) | | | (0.003) | (0.005) |
| *Fixed-effects* | | | | | | | | | | | |
| Player | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Quarter | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Opponent Player | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | | | | | | | | |
| Observations | 45,806 | 46,454 | 46,454 | 46,453 | 46,335 | 46,334 | 46,334 | 45,762 | 45,762 | 45,761 | 45,764 |
| Squared Correlation | 0.211 | 0.349 | 0.081 | 0.152 | 0.214 | 0.213 | 0.214 | 0.212 | 0.211 | 0.212 | 0.213 |
| Pseudo $R^2$ | 0.171 | 0.145 | 0.102 | 0.030 | 0.173 | 0.172 | 0.173 | 0.172 | 0.171 | 0.172 | 0.172 |
| BIC | 72,029.429 | 140,067.468 | 56,742.687 | 273,520.002 | 72,519.107 | 72,577.867 | 72,521.457 | 71,487.035 | 71,549.677 | 71,484.283 | 71,506.228 |

*Clustered standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

**Table 8:** Average difference in winning probability by age: Heterogeneity by the number of moves

| Dependent Variable: | Diff from APG | | | | | |
|---|---|---|---|---|---|---|
| Moves: | 1-30 | 1-60 | 1-90 | 1-120 | 1-150 | 1-180 |
| Model: | (1) | (2) | (3) | (4) | (5) | (6) |
| *Variables* | | | | | | |
| Rank (-) | -2.501*** | -2.203*** | -1.491*** | -1.214*** | -0.955*** | -0.944** |
| | (0.292) | (0.282) | (0.300) | (0.306) | (0.347) | (0.382) |
| Rank Diff (p-o) | -1.023*** | -0.647*** | -0.175 | 0.034 | 0.168 | 0.238 |
| | (0.162) | (0.177) | (0.195) | (0.209) | (0.242) | (0.261) |
| White | 0.131*** | 0.124*** | 0.107*** | 0.092*** | 0.069*** | 0.048*** |
| | (0.010) | (0.010) | (0.011) | (0.012) | (0.013) | (0.014) |
| 7.5 komi | -0.041** | -0.003 | 0.004 | 0.026 | 0.035 | 0.037 |
| | (0.019) | (0.019) | (0.020) | (0.023) | (0.026) | (0.032) |
| Young × Post | -0.209*** | -0.183*** | -0.169*** | -0.156*** | -0.131*** | -0.053 |
| | (0.031) | (0.027) | (0.028) | (0.029) | (0.032) | (0.035) |
| *Fixed-effects* | | | | | | |
| Quarter | Yes | Yes | Yes | Yes | Yes | Yes |
| Player | Yes | Yes | Yes | Yes | Yes | Yes |
| Opponent Player | Yes | Yes | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | | | |
| Observations | 46,454 | 46,438 | 46,284 | 45,310 | 41,492 | 33,510 |
| $R^2$ | 0.349 | 0.267 | 0.210 | 0.168 | 0.141 | 0.139 |
| Within $R^2$ | 0.014 | 0.010 | 0.006 | 0.004 | 0.003 | 0.001 |

*Clustered standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

# APPENDICES

## A. LEELA AND LEELA ZERO

Leela, an AI-powered (APG) Go program, was released in February 2017; it included a stable, deep learning version using Monte Carlo Tree Search (MCTS) and deep learning. It was developed based on the AlphaGo algorithm of Google DeepMind (Silver et al., 2016).[i] Leela is the first APG to surpass human professional players and to be made publicly available on personal computers. A successor with a stronger open-source Go engine, Leela Zero, was released in October 2017 following the publication of an AlphaGo Zero research article by Google DeepMind (Silver et al., 2017). Unlike the original Leela, which uses human knowledge and heuristics in learning, Leela Zero uses only basic rules during self-training. Leela Zero is a faithful reimplementation of the famous Go engine, AlphaGo Zero, and has been made publicly available. We used Leela Zero to evaluate the moves of the players.

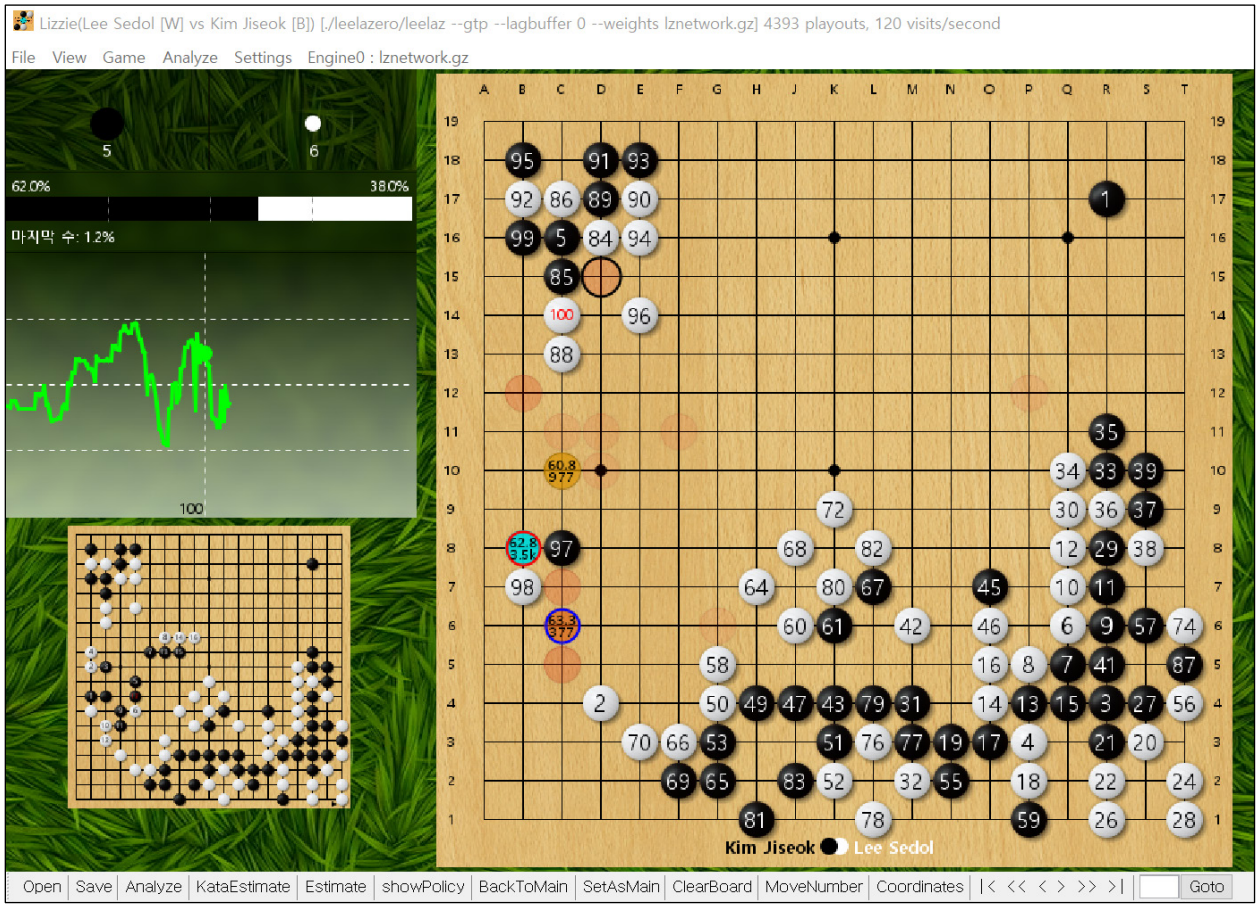
Instead of training a Go engine using expensive Google TPUs (Tensor Processing Units), Leela Zero adopts crowd-sourcing infrastructure using GPUs (Graphics Processing Units) via the OpenCL (Open Computing Language) library.[ii] Leela Zero users can contribute their GPU resources to make Leela Zero stronger. Because of this crowdsourcing training, Leela Zero has rapidly improved over time and continues to improve. Leela Zero provides various Go analysis functionalities but these are not meant to be used directly. Several graphical user interface software programs support Leela Zero so that end-users may utilize various functionalities without hassle. Examples of these interfaces include Lizzie, Sabaki, and GoReviewPartner.[iii]

Leela Zero provides an in-depth analysis of the game, including recommendations for next moves. We visualize what Leela Zero provides for Go analysis using the Lizzie graphical user interface (GUI). Figure A.1 shows a recent Go match by the world champion, Lee Sedol. On the main board, the number on each stone shows the order in which that stone was placed on the board. After the opponent player makes the 100$^{th}$ move (the white stone on C14), Leela Zero recommends multiple moves for the focal player based on MCTS simulations. The cyan-colored point

[i] While AlphaGo and AlphaGo Zero proved the power of AI in Go games, they are not open-source software, nor are user-friendly interfaces provided.
[ii] For more information on OpenCL, please refer to: https://www.khronos.org/opencl/
[iii] These interfaces are available at: https://github.com/featurecat/lizzie/releases (Lizzie), https://sabaki.yichuanshen.de/ (Sabaki), and https://github.com/pnprog/goreviewpartner (GoReviewPartner).

represents the move with the highest recommendation, which has a winning probability of 62.8 percent; this probability is evaluated with 3.5k simulations using MCTS.

**Figure A.1:** Leela Zero and Its Graphical User Interface (GUI)




*Notes*. This is a game between Sedol Lee (white) and Jiseok Kim (black) on 07/26/2019.

In the lower left-hand corner of the screen, Leela Zero provides future simulations of the game, the next sixteen predicted moves. In the upper left-hand corner are shown the current winning probability (62.0 percent) and its change over time (the green line). The graph helps the user evaluate the status of the game and analyze how each move changes the winning probability.

## B. IMPLEMENTATION DETAILS

We used the official version of Leela Zero to analyze the collected Go games. Since Leela Zero improves over time, we fixed the Leela Zero model trained on May 23, 2020 for analysis. We

worked with the GoReviewPartner program to analyze a batch of games. We first analyzed each SGF-formatted file using Leela Zero and saved it into an RSGF-formatted file with Leela_zero_analysis.py code. Then each RSGF file is converted to a CSV format file using r2csv.py code for analysis.

We set five seconds as the time budget for Leela Zero to analyze the winning probability of each move. The five-second time budget is the same setting used in the AlphaGo Zero paper (Silver et al. 2017) to analyze the relative performance among AI Go engines. We ran Leela Zero on a Linux system with four Nvidia Titan-X GPUs and an Inter Core i7-6800K CPU. Each game analysis took approximately twenty minutes; with a single GPU, it would have taken 345 full days to analyze all 25,033 games. We finished our game analysis in about three months by running two to eight GPUs in parallel.

## APPENDIX REFERENCES


Silver D, Huang A, Maddison CJ, Guez A, Sifre L, van den Driessche G, Schrittwieser J, et al. (2016) Mastering the game of Go with deep neural networks and tree search. *Nature* 529(7587): 484–489.

Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, Hubert T, et al. (2017) Mastering the game of Go without human knowledge. *Nature* 550(7676): 354–359.

## C. ADDITIONAL FIGURES AND TABLES

**Figure C.1:** Average difference in winning probability by age: Heterogeneity by the number of moves

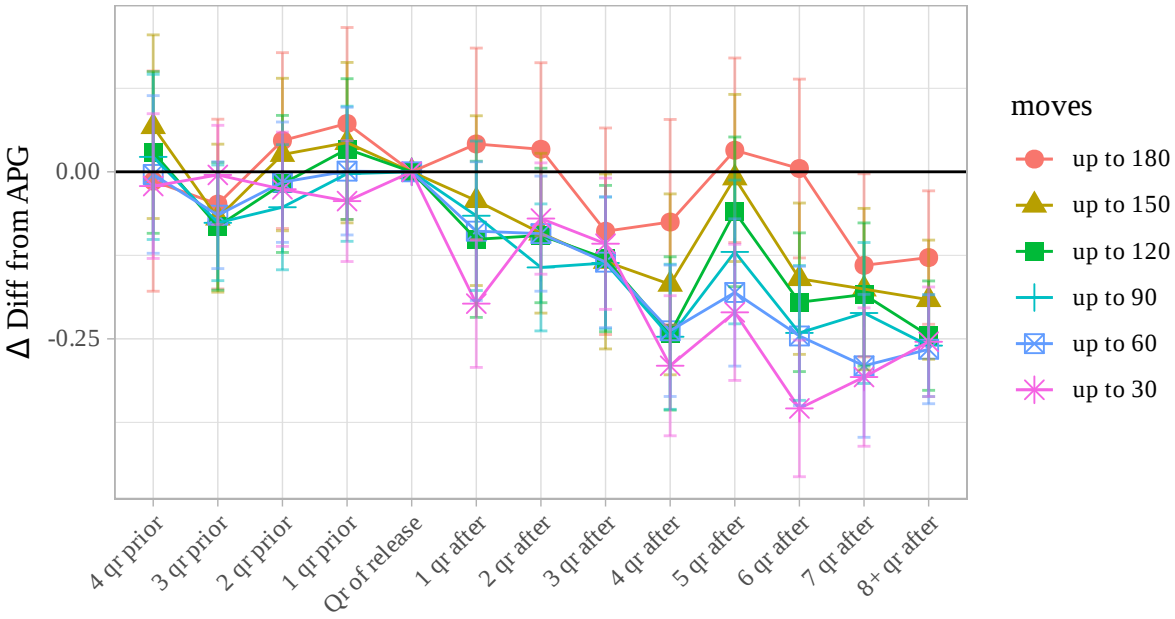


**Table C.1:** Robustness check: Average age as a cutoff for junior group

| Dependent Variable: | Diff from APG | | | |
|---|---|---|---|---|
| Model: | (1) | (2) | (3) | (4) |
| *Variables* | | | | |
| Young | -0.142*** | 0.057** | | |
| | (0.023) | (0.024) | | |
| Rank (-) | -0.783*** | -0.750*** | -1.608*** | -2.396*** |
| | (0.036) | (0.036) | (0.227) | (0.284) |
| Rank Diff (p-o) | -0.124*** | -0.110*** | -0.058** | -0.908*** |
| | (0.027) | (0.027) | (0.024) | (0.163) |
| White | 0.133*** | 0.133*** | 0.131*** | 0.131*** |
| | (0.010) | (0.010) | (0.009) | (0.010) |
| 7.5 komi | -0.027* | -0.026* | -0.026 | -0.043** |
| | (0.016) | (0.016) | (0.016) | (0.019) |
| Young × Post | | -0.370*** | -0.316*** | -0.289*** |
| | | (0.030) | (0.032) | (0.032) |
| *Fixed-effects* | | | | |
| Quarter | Yes | Yes | Yes | Yes |
| Player | | | Yes | Yes |
| Opponent Player | | | | Yes |
| *Fit statistics* | | | | |
| Observations | 46,454 | 46,454 | 46,454 | 46,454 |
| $R^2$ | 0.276 | 0.281 | 0.325 | 0.350 |
| Within $R^2$ | 0.065 | 0.072 | 0.015 | 0.015 |

*Clustered standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

**Table C.2:** Robustness check with three age categories: Young, Mid, and Old

| Dependent Variable: | Diff from APG | | | |
|---|---|---|---|---|
| Model: | (1) | (2) | (3) | (4) |
| *Variables* | | | | |
| Mid Age | -0.095*** | 0.045** | | |
| | (0.022) | (0.022) | | |
| Young | -0.142*** | 0.078*** | | |
| | (0.025) | (0.026) | | |
| Rank (-) | -0.804*** | -0.781*** | -1.503*** | -2.340*** |
| | (0.036) | (0.036) | (0.243) | (0.296) |
| Rank Diff (p-o) | -0.132*** | -0.119*** | -0.059** | -0.962*** |
| | (0.027) | (0.027) | (0.024) | (0.162) |
| White | 0.133*** | 0.134*** | 0.131*** | 0.131*** |
| | (0.010) | (0.010) | (0.009) | (0.009) |
| 7.5 komi | -0.026 | -0.026 | -0.025 | -0.043** |
| | (0.016) | (0.016) | (0.016) | (0.019) |
| Mid Age × Post | | -0.265*** | -0.221*** | -0.198*** |
| | | (0.032) | (0.033) | (0.033) |
| Young × Post | | -0.369*** | -0.328*** | -0.302*** |
| | | (0.034) | (0.040) | (0.040) |
| *Fixed-effects* | | | | |
| Quarter | Yes | Yes | Yes | Yes |
| Player | | | Yes | Yes |
| Opponent Player | | | | Yes |
| *Fit statistics* | | | | |
| Observations | 46,454 | 46,454 | 46,454 | 46,454 |
| $R^2$ | 0.276 | 0.280 | 0.325 | 0.350 |
| Within $R^2$ | 0.065 | 0.070 | 0.014 | 0.014 |

*Clustered standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

**Table C.3:** Placebo test: Random reassignment of the age group

| Dependent Variable: | Diff from APG | | | |
|---|---|---|---|---|
| Model: | (1) | (2) | (3) | (4) |
| *Variables* | | | | |
| Young | -0.003 | -0.008 | | |
| | (0.018) | (0.020) | | |
| Rank (-) | -0.856*** | -0.856*** | -2.364*** | -3.297*** |
| | (0.036) | (0.036) | (0.229) | (0.276) |
| Rank Diff (p-o) | -0.142*** | -0.142*** | -0.068*** | -1.170*** |
| | (0.027) | (0.027) | (0.024) | (0.163) |
| White | 0.133*** | 0.133*** | 0.131*** | 0.131*** |
| | (0.010) | (0.010) | (0.009) | (0.009) |
| 7.5 komi | -0.049*** | -0.049*** | -0.024 | -0.042** |
| | (0.017) | (0.017) | (0.016) | (0.019) |
| Young × Post | | 0.010 | -0.002 | -0.004 |
| | | (0.029) | (0.029) | (0.028) |
| *Fixed-effects* | | | | |
| Quarter | Yes | Yes | Yes | Yes |
| Player | | | Yes | Yes |
| Opponent Player | | | | Yes |
| *Fit statistics* | | | | |
| Observations | 46,466 | 46,466 | 46,466 | 46,466 |
| $R^2$ | 0.274 | 0.274 | 0.322 | 0.347 |
| Within $R^2$ | 0.062 | 0.062 | 0.010 | 0.011 |

*Clustered standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

**Table C.4:** Robustness check: Monthly fixed effect

| Dependent Variable: | Diff from APG | | | |
|---|---|---|---|---|
| Model: | (1) | (2) | (3) | (4) |
| *Variables* | | | | |
| Young | -0.104*** | 0.045** | | |
| | (0.020) | (0.021) | | |
| Rank (-) | -0.819*** | -0.801*** | -1.631*** | -2.490*** |
| | (0.037) | (0.037) | (0.238) | (0.291) |
| Rank Diff (p-o) | -0.138*** | -0.129*** | -0.065*** | -1.013*** |
| | (0.027) | (0.027) | (0.024) | (0.162) |
| White | 0.134*** | 0.133*** | 0.131*** | 0.131*** |
| | (0.010) | (0.010) | (0.009) | (0.010) |
| 7.5 komi | -0.022 | -0.019 | -0.014 | -0.028 |
| | (0.017) | (0.016) | (0.016) | (0.019) |
| Young × Post | | -0.269*** | -0.223*** | -0.206*** |
| | | (0.028) | (0.031) | (0.031) |
| *Fixed-effects* | | | | |
| Month | Yes | Yes | Yes | Yes |
| Player | | | Yes | Yes |
| Opponent Player | | | | Yes |
| *Fit statistics* | | | | |
| Observations | 46,454 | 46,454 | 46,454 | 46,454 |
| $R^2$ | 0.278 | 0.282 | 0.327 | 0.352 |
| Within $R^2$ | 0.064 | 0.068 | 0.013 | 0.014 |

*Clustered standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

**Table C.5:** Robustness check: Alternative definitions of early moves

| Dependent Variable: | Diff from APG | | | | | |
|---|---|---|---|---|---|---|
| | First 40 moves (1-40) | | First 50 moves (1-50) | | First 60 moves (1-60) | |
| Model: | (1) | (2) | (3) | (4) | (5) | (6) |
| *Variables* | | | | | | |
| Rank (-) | -1.823*** | -2.755*** | -1.703*** | -2.514*** | -1.588*** | -2.203*** |
| | (0.224) | (0.277) | (0.239) | (0.296) | (0.232) | (0.282) |
| Rank Diff (p-o) | -0.041 | -1.048*** | -0.035 | -0.894*** | -0.013 | -0.647*** |
| | (0.027) | (0.172) | (0.028) | (0.181) | (0.029) | (0.177) |
| White | 0.128*** | 0.128*** | 0.124*** | 0.125*** | 0.123*** | 0.124*** |
| | (0.010) | (0.010) | (0.010) | (0.010) | (0.010) | (0.010) |
| 7.5 komi | -0.023 | -0.033* | -0.013 | -0.017 | -0.000 | -0.003 |
| | (0.016) | (0.019) | (0.015) | (0.019) | (0.016) | (0.019) |
| Young × Post | -0.220*** | -0.204*** | -0.214*** | -0.202*** | -0.195*** | -0.183*** |
| | (0.029) | (0.029) | (0.028) | (0.028) | (0.028) | (0.027) |
| *Fixed-effects* | | | | | | |
| Quarter | Yes | Yes | Yes | Yes | Yes | Yes |
| Player | Yes | Yes | Yes | Yes | Yes | Yes |
| Opponent Player | | Yes | | Yes | | Yes |
| *Fit statistics* | | | | | | |
| Observations | 46,452 | 46,452 | 46,448 | 46,448 | 46,438 | 46,438 |
| $R^2$ | 0.295 | 0.320 | 0.269 | 0.294 | 0.242 | 0.267 |
| Within $R^2$ | 0.012 | 0.013 | 0.011 | 0.012 | 0.010 | 0.010 |

*Clustered standard-errors in parentheses*
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*