UTMD Working Paper
The University of Tokyo
Market Design Center

UTMD-084

Using Big Data and Machine Learning to Uncover How Players Choose Mixed Strategies

Toshihiko Hirasawa
University of California, Los Angeles

Michihiro Kandori
The University of Tokyo

Akira Matsushita*
Kyoto University

The first version: September 26, 2022
This version: March 28, 2025

Using Big Data and Machine Learning to Uncover How Players Choose Mixed Strategies

By Toshihiko Hirasawa, Michihiro Kandori, and Akira Matsushita*

Draft: March 28, 2025

We examined how humans learn to choose mixed strategies using our unique big experimental dataset with approximately 75,000 observations. We compared the out-of-sample predictive power of conventional behavioral models and machine learning models and found that a version of the deep learning model (LSTM) substantially outperforms the conventional models. The superiority of the machine learning model is noticeable only when the data size is an order of magnitude larger than that of the typical lab dataset. We tried to open the black box of LSTM and obtained an improved behavioral model with nearly equal predictive power. We provide several key steps one can follow to improve existing behavioral models by means of machine learning and big data.

The purpose of this paper is two-fold. On the methodological side, our work provides a list of key steps one can follow to improve existing empirical models by means of big data and machine learning. On the conceptual side, we try to uncover how human players learn to behave when one would like to make one's own behavior unpredictable (i.e., how they learn to play mixed strategies). Using a unique experimental dataset we collected, which is among the largest for a single-treatment lab experiment, we compared the conventional behavioral models and several leading machine learning models, and we found that the machine learning models performed substantially better. One of the most important findings we obtained is that the superiority of machine learning models is noticeable only when the dataset size is an order of magnitude larger than that of the typical lab dataset. The superiority of the machine learning models shows that the existing

* Hirasawa: Department of Economics, UCLA, Bunche Hall 315 Portola Plaza, Los Angeles, CA, 90095 (email: toshi224hirasawa@gmail.com). Kandori: Department of Economics, University of Tokyo, 7-3-1 Hongo Bunkyo-ku 113-8654 Tokyo Japan (email: kandori@e.u-tokyo.ac.jp). Matsushita: Department of Informatics, Kyoto University, Yoshidahonmachi Sakyo-ku Kyoto-shi 606-8501 Kyoto Japan, and UTMD (email: amatsushita@i.kyoto-u.ac.jp). This is a substantially revised version of the earlier version under the same title. We would like to thank Annie Liang and Tetsuya Kaji for their helpful suggestions and discussion. We are also grateful to the seminar audiences at Northwestern, Harvard-MIT, Chicago, Caltech, NYU Abu Dhabi, Japanese Economic Association, UCLA, Peking, Tokyo, Tohoku, Osaka, and Kyoto for their helpful comments and discussion. We thank Nanami Aoi, Shinji Koiso, and Ayano Yago for their research assistance. The first author acknowledges financial support from the Yoshida Scholarship Foundation. The second author acknowledges financial support under JSPS KAKENHI Grants JP20H00609, JP21K01399, and JP21H04979.

models are missing out on some important regularities of human behavior, and it also shows to what extent the existing models can possibly be improved. We, therefore, attempted to open the black box of our machine learning models, and we ultimately obtained a new behavioral model that has almost the same predictive power as our best machine learning model.

The conceptual research question we address is how humans behave when (i) one needs to make one's own actions unpredictable and (ii) at the same time, one needs to predict what actions an opponent will take. Examples include tax auditing, terrorist attacks vs. airport security guards (Tambe, 2011), tennis serves (Walker and Wooders, 2001; Gauriot, Page and Wooders, 2023), and penalty kicks in soccer games (Palacios-Huerta, 2003; Chiappori, Levitt and Groseclose, 2002). Such a situation can be formalized as a game with a unique mixed strategy equilibrium. Previous research (e.g., the papers cited above as well as O'Neill (1987) and Camerer (2011)) has revealed that mixed strategy equilibria describe human behavior in field and lab data reasonably well, while it has also been shown that humans do not exactly follow a mixed strategy equilibrium (e.g., Brown and Rosenthal, 1990). If we are called upon to play a game with a unique mixed strategy equilibrium, rather than calculating and following the mixed strategy equilibrium, we use our intuition, hunches, and some kind of limited reasoning. It is not yet fully understood exactly what kind of cognitive processes are employed, and we aim to answer this question by using unique big data and machine learning models.

Our dataset examines one of the simplest possible games with a unique *non-trivial* mixed strategy equilibrium. It is a card game invented by O'Neill (1987), where each player, the "red" and the "black" player, has four cards, K, 1, 2, and 3, and chooses one of those cards at the same time as the opponent. The winner is determined by a rather complicated set of rules. As a result, unlike with rock-paper-scissors or the matching pennies game, the equilibrium probability distribution over the four cards is *not* uniform, and the black player has a higher chance of winning in the equilibrium. Therefore, the subjects need to think carefully to figure out what to do, rather than just naively mixing *uniformly.* Such a situation is ideal for addressing our conceptual research question. The dataset was collected in a Coursera online course on game theory and contains data on more than 5,000 participants. Each player played the game 30 times with the same opponent, providing roughly 75,000 observations (for each player's role). Subsection II.B shows that this is one of the largest datasets for a single treatment in economic laboratory experiments. Note that our dataset shows how human players *learn* to play this game over the 30 rounds, where the learning rule maps the past history of play to the choice probability of current actions. This learning rule is the focus of our study.

To uncover how human players mix, we tried to improve conventional behavioral models with the help of the machine learning approach. The machine learning approach is different from conventional econometrics, at least in two respects.

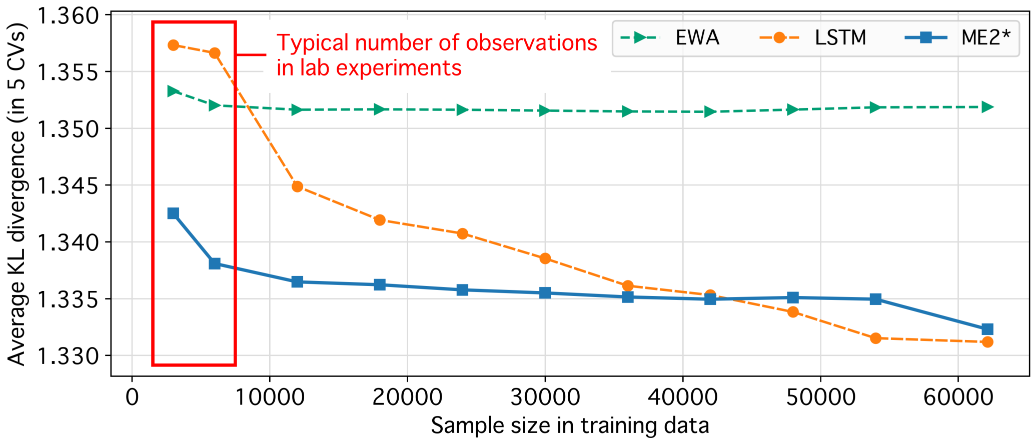
First, it offers functional forms that are new to economists. Second, and more importantly, it evaluates models by their *external validity*: Models are evaluated by their "out-of-sample" prediction errors in a dataset that is *not* used for parameter estimation. External validity is a hallmark of established models in hard sciences, but it has not been the main focus in economics, presumably because available datasets are usually not big enough for systematically checking external validity. Given our unique large dataset, however, we were able to compare the external validity of conventional behavioral models, such as a leading model of learning in games, EWA (Camerer and Ho, 1999), and the serial correlation model, and leading machine learning models, including the decision tree, LightGBM, LASSO models, and deep learning models (DNN and LSTM).

We found that LSTM, which was used for the earlier version of Google Translate, substantially outperformed the conventional behavioral models. LSTM is a type of deep learning model. Such models have a large number of parameters (more than 5,000), the interpretation of which is not immediately clear. Rather than directly opening the black box of LSTM, we tried to gain insights from more interpretable machine learning models, namely the decision tree and LASSO models. Those models suggest that the subjects in our data are subject to a particular form of serial correlation. By incorporating the serial correlation and the fact that the opponent is also subject to the same tendency into the EWA model, we obtained a new, improved model that captures how humans learn to choose mixed strategies, which we call *mixing EWA*. It captures almost all the predictive power of our best machine learning model (LSTM).

Moreover, we artificially reduced the size of our dataset used for parameter estimation to examine how the out-of-sample prediction errors of the conventional behavioral models and the machine learning models would change. Figure 1 shows the results. Under the typical lab dataset size with at most a few hundred subjects, the superiority of the machine learning model is not noticeable. As the dataset size increases, however, the performance of machine learning keeps on improving, while the conventional model (EWA) does not show any noticeable improvement. These results suggest that to clearly detect the limitations of conventional models using machine learning models, we need a dataset that is an order of magnitude larger than the typical lab dataset. We view this as one of the main findings of this paper.

Why does the performance of the machine learning model keep on improving as the dataset size increases? This is because the machine learning model has a large number of parameters (more than 5,000). One might argue that it is not surprising that a model with a larger number of parameters fares better, but this view is misplaced. That would certainly be true if we compared the *goodness of fit* of models, but it is by no means true when we compare the external validities of models. If the true model has only one parameter, for example, the out-of-sample prediction error of this simple model should be better than more complicated models. This is because models with redundant parameters "overfit" the data

FIGURE 1. OUT-OF-SAMPLE PREDICTION ERRORS WITH DIFFERENT SAMPLE SIZES IN TRAINING DATA



Note: Out-of-sample prediction errors of the traditional model (EWA), our best machine learning model (LSTM), and our novel mixing EWA model (ME2*) for the red player with different sample sizes in training data. The methodology for performance measurement is detailed in Section V. The range of observations typical in laboratory experiments, usually fewer than 10,000 as shown in Figure 2, is highlighted with a red rectangle.

used for parameter estimation (the training data) and fare worse in the test data. The superiority of our best machine learning model in terms of out-of-sample prediction error shows that the model does capture regularities of human behavior that are common both in the training data and the test data.

Figure 1 further illustrates that our best mixing model, a version of Mixing EWA that we call ME2*, significantly outperforms the traditional EWA model, even when tested on the typical dataset size found in lab experiments. This improvement is achieved because our model incorporates patterns in human behavior identified by the machine learning model despite using substantially fewer parameters ($\simeq 100$). By decoding the regularities captured by the machine learning model trained on large datasets, we developed a compact model that is applicable to a broader range of constant-sum games beyond O'Neill's game.

On the technical side, we propose five key steps one can follow, including relevant statistical tests, to improve the existing behavioral models by means of machine learning and big data. We term this procedure Detect, Capture, and Decode (DCD). A summary description of this procedure can be found in the concluding section. This paper, in which we were able to improve the conventional model to achieve the same level of predictive power as the best machine learning model, provides an instance of a successful application of this procedure.

A. *Related Literature*

An overview of laboratory experiments on games with a unique mixed strategy equilibrium can be found in Chapter 3 in Camerer (2011). In the case of the specific game we analyzed, while O'Neill (1987) stresses the accuracy of the mixed strategy equilibrium predictions, Brown and Rosenthal (1990) reexamined O'Neill's result and showed that the subjects were not playing an i.i.d. mixture over time.

Our paper is also related to the literature on learning in games. See Fudenberg and Levine (1998), Chapter 13 in Camerer, Loewenstein and Rabin (2004), and Chapter 6 in Camerer (2011) for surveys. Leading models include reinforcement learning (e.g., Roth and Erev (1995) and Erev and Roth (1998)) and belief learning such as fictitious play (Brown, 1951). Camerer and Ho (1999) propose an influential generalized model incorporating both reinforcement learning and belief learning, called the experience-weighted attraction (EWA) model, which we employed as a leading learning model in games in this paper.

Finally, our paper contributes to the rapidly growing literature on machine learning and economics, overviews of which can be found in Athey and Imbens (2019), Mullainathan and Spiess (2017), and Camerer (2019). Our paper is directly related to the previous studies about improving existing behavioral models using machine learning, which include Peysakhovich and Naecker (2017), Peterson et al. (2021), and Fudenberg and Liang (2019). The first two papers concern single-person decision problems about lotteries, while the last one analyzes the initial play in simple games given by payoff matrices. The challenge in those papers is to find good non-linear *functional forms* to predict subjects' behavior $y = f(x)$, where x is a vector of *a priori* given, a small number of continuous variables, such as the probabilities and prizes of lotteries, or entries of payoff matrices. In contrast, our paper suggests another way in which machine learning can help: the selection of the right-hand side variable x. In our study, y represents the choice probabilities of current actions, and x should represent relevant aspects of the past history of play. Even if we focused on 2-period histories, the number of candidate variables (dummies for subsets of history) would be $2^{256}-1$. The challenge we face is to figure out the relevant variables from an astronomical number of candidates.

Our paper, together with the works discussed above, suggests that the dataset size should be an order of magnitude larger than the typical lab dataset when trying to improve existing behavioral models using machine learning. Peysakhovich and Naecker (2017), one of the earliest papers that tried to improve existing models using machine learning, found that there was little room for improvement. Their study is based on data collected from 300 subjects. Peterson et al. (2021), Fudenberg and Liang (2019), and our paper, in contrast, involve 14,711, 6,887, and 5,178 subjects, respectively, with the finding that existing behavioral models can be substantially improved with the help of the machine learning approach.

Finally, Fudenberg and Liang (2020) and Fudenberg et al. (2021) explore the optimal trade-off between the accuracy and simplicity of models. They introduce a measure of accuracy called *completeness*, and we utilize a version of completeness to compare the various models we consider.

Table 1—Payoff matrix of the O'Neill's stage game

		Black			
		1	2	3	K
Red	1	0, 1	1, 0	1, 0	0, 1
	2	1, 0	0, 1	1, 0	0, 1
	3	1, 0	1, 0	0, 1	0, 1
	K	0, 1	0, 1	0, 1	1, 0

I. O'Neill's Game

In this paper, we consider *O'Neill's game* as introduced by O'Neill (1987). It is a normal-form game with two players, a red (R) player and a black (B) player. Each player $i \in I \equiv \{\text{R, B}\}$ chooses one of four cards, $a_i \in C \equiv \{1, 2, 3, \text{K}\}$ (Ace, Two, Three, and King), as their action. The payoff matrix is shown in Table 1.[1] In words, the red player wins if and only if (1) both players choose K, or (2) both players choose numbers (1, 2, or 3) and those two numbers are different. The black player wins in the remaining cases. This game was designed to be the simplest possible one with a non-trivial mixed strategy equilibrium.[2]

O'Neill's game has a unique mixed strategy Nash equilibrium: both players play K with a probability of 0.4 and play 1, 2, and 3 with a probability of 0.2, respectively. In the equilibrium, the red player has a lower chance of winning (probability 0.4). Hence, unlike in the rock-paper-scissors game, the mixed strategy is not trivial. As a result, our subjects need to think carefully about what to do rather than just naively choose all actions uniformly, which makes O'Neill's game ideal for examining how human players mix.

II. Data

A. Data Collection

We collected our dataset in an online introductory game theory course offered by one of the authors.[3] In the first week, before taking lectures on Nash equilibrium and mixed strategy equilibrium, the students were asked to play 30 rounds of this game with someone else and submit the results on the course web page. The

[1]In O'Neill's experiment, Joker was used instead of K, and a winner (loser) gets (loses) a payoff of 5.

[2]More precisely, O'Neill (1987) shows that the game is the unique normal form game that satisfies the following conditions: (i) There are binary payoffs (win or lose) for each player so that the degree of risk aversion does not change the mixed strategy equilibrium, (ii) neither player has two identical strategies, (iii) neither player has a dominant strategy, and (iv) the game is not completely symmetrical in strategies. Any other game satisfying the conditions above has at least as many strategies for each player.

[3]Coursera course "Welcome to Game Theory," by Michihiro Kandori, an introductory course that assumes no background knowledge of mathematics or economics (https://www.coursera.org/learn/game-theory-introduction). The course started in February 2015 and is still available.

students were told, "Your participation in this activity should be on a completely voluntary basis, and your answers will have no bearing on your course grades."

Hence, our dataset is different from the typical lab dataset in the following respects. First, the experiments were unsupervised. We asked the students to find a partner, play the game, record the results, and report back. Second, we do not know the identities of the students' partners, who we expect to be their friends or family members. Third, no financial rewards were given to the subjects. We do not think these features pose a serious problem for the following reasons. First, we asked the students to play this game with a deck of cards. People are naturally motivated to win when playing a card game, even if no monetary payment is made. Second, we clarified that participation has nothing to do with the course grade, while participation is costly in terms of time and effort. This fact likely discouraged attempts to submit fake data. A person who submits fake data incurs the time and effort of reading and understanding the instructions, cooking up data for 30 rounds, and uploading the file, even though no reward is given.

Our dataset here contains 2,781 pairs (5,562 participants) of play in total, submitted from February 2015 to April 2022. However, 192 pairs of data in which a player's action was missing in at least one period were eliminated, and we used the remaining 2,589 data pairs (5,178 participants).

B. The Uniqueness of Our Big Data

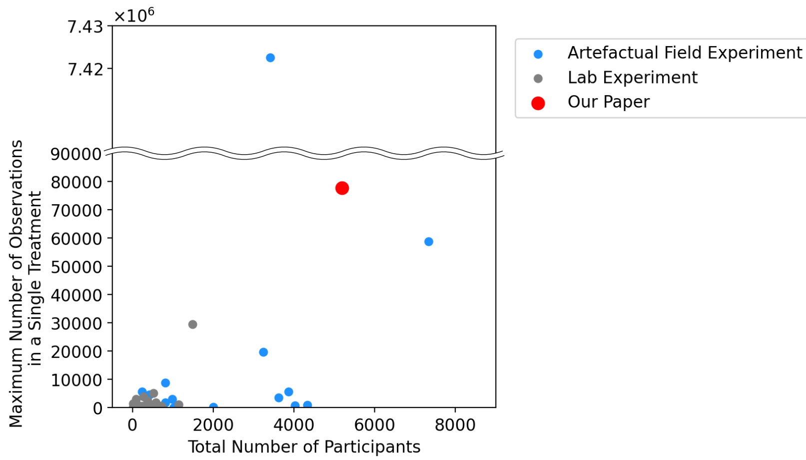
The uniqueness of our dataset lies in the number of participants, especially the number of observations. We collected data from 2,589 pairs (5,178 participants). Each pair played the O'Neill's stage game for 30 rounds. Hence, the total number of observations is $77,670 = 2,589 \times 30$ for the red player and the black player, respectively.

The number of observations in our dataset is one of the largest for a single treatment. Figure 2 plots the total number of participants and the maximum number of observations in a single treatment in all experimental papers published in the top 5 journals between 2010 and 2021.[4] The number of observations in our dataset is greater than those in all those papers except for one, whose largest number of observations concerns "market average probability assessments" backed up by the high-frequency market data of a prediction market in the UK (Augenblick and Rabin, 2021). Although there is another paper in Figure 2 whose maximum number of observations is close to ours, it is clear that the number of observations in our paper is one of the largest.

The uniqueness of our dataset is especially clear in lab experiments. On average, each experiment has 438.111 participants with a standard deviation of

[4]The top 5 journals are American Economic Reviews, Econometrica, Quarterly Journal of Economics, Journal of Political Economy, and Review of Economic Studies. All experimental papers published in those journals are compiled in Nunnari, Congiu and Emiliano (2022).

Figure 2. Total number of participants and maximum number of observations in experiments published in the top five journals between 2020-2021



Note: The total number of participants and the maximum number of observations in a single treatment in each paper published in 2020 and 2021. There are 39 papers in total in those two years. Of the 39 papers, 18 are classified as lab experiments, and 21 are classified as artefactual field experiments. Artefactual field experiments are defined by Harrison and List (2004), and this category includes experiments conducted by Amazon Mechanical Turk (MTurk).

402.392, and each treatment has $2,975.944$ observations with a standard deviation of $6,772.554$.

C. Descriptive Statistics

This section briefly describes some features of our 2,589 pairs $\times$ 30 periods of data based on summary statistics. First, we observe that the choice frequencies of cards and win rates are largely consistent with the Nash equilibrium prediction at the aggregate level, which replicates what O'Neill found (Table 2).

However, the time series of the subjects' choices cast doubt on the hypothesis that the subjects play the mixed Nash equilibrium independently every period. Figure 3 illustrates the trajectories of the average choice frequencies of all the red and black players in each period. The variations in the aggregate frequencies of the four cards over time are rather pronounced, which is not what one would expect from i.i.d. sequences with 2,589 observations in each period.

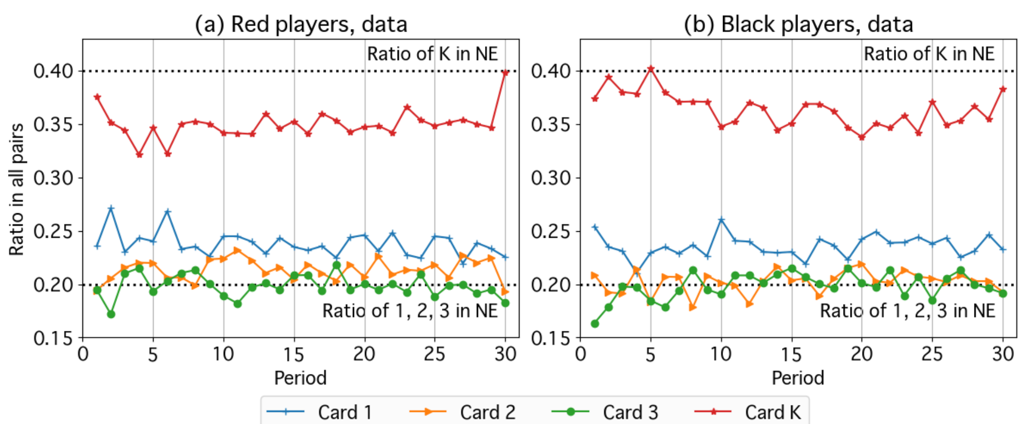
To examine this point more formally, Table 3 shows how the card frequencies depend on the previous action of the focal player. The rows in this table should be similar if the subjects played the i.i.d. mixed strategy, but they are rather different. This shows that the choice of current action depends at least on the previous action, a violation of i.i.d. choice. In fact, Brown and Rosenthal (1990) rigorously showed that the null hypothesis that the subjects in O'Neill's data

TABLE 2—ACTION DISTRIBUTIONS IN OUR DATA AND IN O'NEILL'S DATA

	(1) Action distribution								(2) Win rates	
	Red				Black				Red	Black
	K	1	2	3	K	1	2	3		
NE	0.400	0.200	0.200	0.200	0.400	0.200	0.200	0.200	0.400	0.600
Our data	0.350	0.238	0.214	0.198	0.364	0.236	0.202	0.198	0.419	0.581
O'Neill	0.362	0.221	0.215	0.203	0.426	0.226	0.179	0.169	0.409	0.591

Note: Action distributions in our data and in O'Neill's data (summarized in Brown and Rosenthal (1990)). O'Neill collected 2625 observations of the O'Neill's stage game results (25 pairs × 105 periods).

FIGURE 3. PERIOD TRANSITION OF AVERAGE CHOICE PROBABILITIES



played the i.i.d. mixture can be statistically rejected. What mechanisms guide the subjects' behavior? This is the question we address in this paper.

III. Conventional Models

The behavior of our subjects may be explained by models proposed in the existing literature, which we call conventional models, and we will examine if machine learning models can improve those models. The conventional models include a simple benchmark model with only a constant term (naive i.i.d. model), models that capture the fact that humans are not good at generating i.i.d. sequences and tend to create negative correlation (serial correlation models), and the EWA model, a leading behavioral model of learning in games. The machine learning models we consider are LASSO, decision tree, LightGBM, and deep learning models such as DNN and LSTM. In this section, we explain the conventional models

TABLE 3—TRANSITION MATRIX OF ACTIONS

(a) Red players

		Action at t: 1	2	3	K
Action at $t-1$	1	0.184	0.241	0.216	0.358
	2	0.264	0.161	0.220	0.355
	3	0.262	0.230	0.148	0.361
	K	0.246	0.219	0.202	0.333
Average		0.238	0.214	0.198	0.350

(b) Black players

		Action at t: 1	2	3	K
Action at $t-1$	1	0.181	0.222	0.215	0.381
	2	0.252	0.155	0.223	0.370
	3	0.260	0.224	0.150	0.366
	K	0.247	0.203	0.203	0.347
Average		0.236	0.202	0.198	0.364

Note: The numbers represent the conditional choice probability of (a) the red player's actions and (b) the black player's actions. Each line indicates the empirical probability distribution over current actions $a_i^t = 1, 2, 3, K$ conditional on a certain previous action a_i^{t-1}.

that we adopted.

Before discussing the models, we first introduce some notation. In every period $t \in T \equiv \{1, 2, \ldots, 30\}$, each player $i \in I = \{R, B\}$ simultaneously chooses one of the four cards, $a_i^t \in C \equiv \{1,\ 2,\ 3,\ \mathrm{K}\}$. We denote the history of action profiles before period t by $h^t \equiv ((a_{\mathrm{R}}^1, a_{\mathrm{B}}^1), (a_{\mathrm{R}}^2, a_{\mathrm{B}}^2), \ldots, (a_{\mathrm{R}}^{t-1}, a_{\mathrm{B}}^{t-1}))$. For notational convenience, we define the initial dummy history $h^0 \equiv \emptyset$.

Each pair in our dataset is indexed by $s = 1, 2, \ldots, S$. For the estimation of the parameters of the conventional models we report in this section, we use all the data we have ($S = 2,589$). In the performance comparison discussed later in this paper, on the other hand, we randomly split our dataset into training data and test data, and estimate model parameters using the training data only. This procedure will be explained in detail in Section IV.

A. *General Formulation and the Challenge of Model Selection*

Any econometric models that explain our data would specify the probability that player i chooses action a in round t given the previous history of play h^t, namely $P(a_i^t = a \mid h^t)$. If each action is chosen with a positive probability after any history in an econometric model, we call it a *full-support* model. We first show that any full support model can be represented by the following logit formula. The probability that a subject in the role of player $i \in I$ chooses action $a \in C$ at period t is given by

$$P_i^a(t) \coloneqq P(a_i^t = a \mid h^t) = \frac{\exp\left\{\beta_{i,a}^{\mathrm{T}} x_i(h^t)\right\}}{\sum_{c \in C} \exp\left\{\beta_{i,c}^{\mathrm{T}} x_i(h^t)\right\}}, \tag{1}$$

where $x_i(h^t)$ is an m-dimensional vector of variables (including the constant term) that depends on the history h^t, where m is the number of right-hand side variables used in the model. In what follows, for simplicity, Equation (1) is expressed as $P_i^a(t) \propto \exp\left\{\beta_{i,a}^{\mathrm{T}} x_i(h^t)\right\}$.

Any full-support model can be written as Equation (1) for the following reason. Since probabilities of actions (cards) add up to 1, the ratios of probabilities (odds), such as $P_i^a(t)/P_i^K(t)$, $a = 1, 2, 3$, uniquely determine the probabilities of actions. Hence, any full-support model can be expressed as

$$\frac{P_i^a(t)}{P_i^K(t)} = g_a(h^t),$$

while Equation (1) boils down to

$$\frac{P_i^a(t)}{P_i^K(t)} = \exp\left\{(\beta_{i,a} - \beta_{i,K})^{\mathrm{T}} x_i(h^t)\right\}. \tag{2}$$

If we let $x_i(h^t)$ be the vector of indicator functions for histories $(\cdots, \mathbf{1}_{\hat{h}^t}, \cdots)$, where $\hat{h}^t$ runs over all possible histories, and denote the element of vector $\beta_{i,a}$ that corresponds to the coefficient of $\mathbf{1}_{\hat{h}^t}$ by $\beta_{i,a,\hat{h}^t}$, then the left-hand side of (2) is simply equal to $\exp\left\{\beta_{i,a,h^t} - \beta_{i,K,h^t}\right\}$. Therefore, by setting $\beta_{i,a,h^t} - \beta_{i,K,h^t} = \log g_a(h^t)$, any full-support model is represented by (1). Equation (2) also shows that $(\beta_{i,a} - \beta_{i,K})$ is identified but $\beta_{i,a}$ and $\beta_{i,K}$ are not.

The conventional models we consider are estimated by the method of maximum likelihood. Our ultimate goal is to find the optimal choice of the right-hand side variables in (1) that provides the minimum out-of-sample prediction error. However, comparing all possible combinations of the right-hand side variables is computationally infeasible. Even when we assume that the player's choice depends only on the history of actions in the last two periods, the number of two-period histories is $(4 \times 4)^2 = 256$, and the number of dummy variables of subsets of two-period histories of play is equal to $2^{256} - 1$, an astronomically large number. Furthermore, the number of feasible combinations of those dummy variables is even larger. It is clearly impossible to compare the out-of-sample prediction power of all those models.

B. *Naive Benchmark and Serial Correlation*

As a benchmark of our analysis, we first estimate the model (1) with only constant terms. The maximum likelihood estimation of this model provides the choice probability (i.i.d. over time) that is equal to the empirical frequencies of cards in the entire dataset. We call this the naive i.i.d. model.

The naive model specifies that the player naively mixes in the same way over time, and it is assumed that actions are serially uncorrelated. However, studies in the field of psychology have shown that humans are poor at generating an i.i.d. random sequence and end up creating negative serial correlation (c.f., Bar-

Hillel and Wagenaar, 1991). The following model incorporates serial correlation based on the past n periods:

$$P_i^a(t) \propto \exp\left(c_i^a + \sum_{s=1}^{n} \alpha_i^{a,s} 1\left\{a_i^{t-s} = a\right\}\right). \tag{3}$$

In this model, the target mixing probability is captured by the constant terms c_i^a, $a = 1, 2, 3, K$, and the probability of player i choosing action a is affected if the same action was chosen in the past n periods.

In Equation (3), although the coefficients $\alpha_i^{a,s}$ can be different across $a = 1, 2, 3, K$, we observed that a model specification with the same coefficients for all number cards yielded superior predictive performance.[5] Therefore, we will adhere to this restricted specification.

We call the model in Equation (3) with that parameter restriction *serial correlation of order n*. We estimated the models with $n = 1, 2, 3$, and 4 and found the model with $n = 4$ was the best. In what follows, we report the models with $n = 1$ and 4.

C. *Experience-Weighted Attraction (EWA) Model*

The models in the previous subsection capture the subjects' attempt to make themselves unpredictable, but those models ignore the strategic thinking inherent in the game they play. The subjects in our data set must have tried to learn which action fares well and what the opponent is going to do. A host of models of *learning in games* have been proposed in the behavioral economics literature. Many of them fall into the following two groups: reinforcement learning (RL) and belief learning (BL).[6] Reinforcement learning is based on the idea that, if an agent chooses an action and the outcome is good, that action is "reinforced" in the sense that the agent is more willing to choose it. This reinforcement process can be formulated as a model in which an agent has some "attraction" to each action, which embodies how much payoff the agent has gained in the past. In a belief learning model, an agent forms a belief about the choice probability of the opponent's action based on the weighted average of the past frequencies of the opponent's actions and plays the best response to the belief.

Camerer and Ho (1999) developed the experience-weighted attraction (EWA) learning model, which incorporates reinforcement learning and belief learning models as its special cases. In their original EWA model, each player $i \in I$

[5] We compare the performance of the restricted and unrestricted Serial Correlation of order 4 models in Appendix C.1

[6] The term "reinforcement learning" has been used to mean somewhat different ideas. We use this term in the sense used in the psychology and neuroscience literature and note that this is not related to the term commonly used in the computer science literature.

chooses action $a \in C$ at period t with probability,

$$P_i^a(t) \propto \exp\Big\{\lambda_i A_i^a(t-1)\Big\},$$

where

$$N_i(t) = \rho_i N_i(t-1) + 1,$$

$$A_i^a(t) = \frac{\phi_i N_i(t-1) A_i^a(t-1) + \big[\delta_i + (1-\delta_i)\mathbb{1}\left\{a_i^t = a\right\}\big]\pi_i(a, a_{-i}^t)}{N_i(t)}, \tag{4}$$

and $\lambda_i \in [0, \infty)$, $\rho_i, \phi_i, \delta_i \in [0, 1]$. Here, $A_i^a(t)$ is interpreted as player i's attraction to action a at period t.

The key parameter is δ_i, which measures how the model combines the ideas of reinforcement learning and belief-based learning. It can be shown that the EWA model with $\delta_i = 0$ is a version of the reinforcement learning model, while $\delta_i = 1$ provides a version of the belief-based learning model.

We estimated the parameters using maximum likelihood estimation with all sample data. Table E13 in Appendix E.1 presents the results. The estimates of the key parameters in EWA, $\delta_{\mathrm{R}} = 0.385$ and $\delta_{\mathrm{B}} = 0.000$, suggest that the red player plays a mixture of RL and BL, while the black player plays the pure RL.

IV. Machine Learning Models

In this section, we provide a brief overview of machine learning in comparison to conventional econometric models. Machine learning refers to a class of models that make predictions or decisions based on observable data. For our purposes, we focus on machine learning models to make a prediction:

$$y = f(x \mid \beta),$$

where x is the input data, y is the predicted outcome, and β is the vector of parameters. This looks exactly the same as the conventional econometric models, but there are some notable differences.

First, machine learning offers new functional forms that have not been employed in conventional econometrics. Leading examples include LASSO, decision trees, LightGBM, and deep learning, which we explain and use in later sections.

Second, while the typical goal of econometric work is to obtain an accurate estimation of a given model's parameters in a given dataset, machine learning research tries to find better models in terms of their out-of-sample prediction errors. The dataset is partitioned into training data and test data, and the parameter values of models are determined using the training data. The models are then compared by their prediction errors in the test data.

As explained above, if we split the dataset only once, we can only perform a prediction contest once. In machine learning research, however, it is common to conduct prediction contests multiple times on the same dataset. This is called cross-validation (CV) and works as follows. First, the whole dataset is randomly split into K subsets of an (almost) equal size, with the typical value of K being 5. Since our dataset contains 2,589 pairs, we created four subsets of 518 pairs and one subset of 517 pairs. Call them $D_1, \ldots, D_5$. Second, the first subset D_1, which accounts for 20% of the whole dataset, is set aside as test data, and parameter fitting is conducted on the remaining 80% of training data $D_2 \cup \cdots \cup D_5$. We then evaluate the performance of the trained model using the test data D_1. In general, we set aside D_k as test data, conduct parameter fitting on the remaining training data $\bigcup_{h \neq k} D_h$, and then measure the model's predictive performance using D_k. This process is repeated for $k = 1, \ldots, 5$, and the models are compared in terms of their *average* prediction errors in the five rounds of cross validation.

In what follows, we compare the external validities of the conventional models and leading machine learning models via cross-validation in our unique big dataset.

A. *Choice of Loss Function*

The machine learning models presented below are trained to minimize the prediction error defined by a loss function. Note that a machine learning model assigns to each set of features X of the past history of play a vector of probabilities $P = (P(1), P(2), P(3), P(K))$, where each $P(c)$ is a predicted probability of card c in the current period. While the prediction of a model is probabilistic, the observed data records realized actions, which can be represented by a degenerated probability distribution $q = (q(1), q(2), q(3), q(K))$, where one of its element is one and the others are zero. The "distance" between the probabilistic prediction P and the realized action q is represented by a loss function $l(P, q)$, and there are many possible specifications. One of the important steps in improving the conventional models by machine learning is the choice of loss function for the machine learning models. We argue that in the present context, the Kullback-Leibler (KL) divergence is the appropriate choice for the reason below.

$$l(P, q) = \sum_{a \in C} q(a) \log\left(\frac{q(a)}{P(a)}\right), \tag{5}$$

where $0 \times \log(0)$ is interpreted as $\lim_{x \downarrow 0} x \log x = 0$. When the realized action is a', this expression boils down to

$$l(P, q) = -\log P(a').$$

Hence, if we have observations $(P_n, q_n), n = 1, ..., N$ in the dataset, the total loss function of a model is

$$\sum_{n=1}^{N} l(P_n, q_n) = -\sum_{n=1}^{N} \sum_{a \in C} q_n(a) \log(P(a)),$$

which is nothing but the log-likelihood of the model. In summary, minimizing the loss defined by the KL divergence is equivalent to the method of maximum likelihood for parameter estimation, which is used for the conventional models we considered and also for the LASSO model described below. In this sense, the KL divergence is the right choice for the loss function used to train our machine learning models because the conventional models and machine learning models are both trained to minimize the same loss function. The same loss function should also be used when we compare the out-of-sample predictive power of those models (Section V).

B. LASSO

LASSO (Least Absolute Shrinkage and Selection Operator) appears to be a promising model for our purpose because it performs an automatic selection of variables in our general model (1). We consider $2,448$ variables (for each card $a \in C$) as potential covariates of the multinomial logit model (1) and add an L1-penalty term into the log-likelihood.[7] That is, the LASSO estimator $\hat{\beta}$ is the solution to the following penalized log-likelihood maximization problem:

$$\max_{\beta} LL(x \mid \beta) - \lambda ||\beta||_1, \tag{6}$$

where LL is the log-likelihood function

$$LL_i(x \mid \beta) \equiv \sum_{s=1}^{S} \sum_{t=1}^{30} \left[\sum_{a \in C} q_a^{s,i,t} \beta_{i,a}^{\mathrm{T}} x_i(h_s^t) - \ln \left(\sum_{a' \in C} \exp\left\{ \beta_{i,a'}^{\mathrm{T}} x_i(h_s^t) \right\} \right) \right]$$

for each $i = \mathrm{R}, \mathrm{B}$. Here, $q_a^{s,i,t}$ is equal to 1 if player i in pair s chooses action a in round t, and otherwise, it is equal to zero, while $x_i(\cdot) \in \mathbb{R}^m$ is the vector of candidate variables whose values are given by the history of play. If we denote the k-th element of the coefficient vector β as β_k, then $||\beta||_1$ is expressed as $\sum_k |\beta_k|$. Since the graph of $|\beta_k|$ has a kink at 0, the maximization problem (6) typically has a "corner solution" where $\beta_k = 0$ for some (and in fact, many) k, which means that the associated variables are excluded from the model.

We employed a total of 2,448 candidate variables. They are mostly dummy variables about what happened in the last 4 periods. In addition, we also in-

[7] The number of parameters we initially have created was 2,702. Dropping variables whose values were exactly the same as another variable then left us with 2,448 variables.

Table 4—Number of Coefficients Selected by LASSO

	$\beta_{\mathrm{R},1}$	$\beta_{\mathrm{R},2}$	$\beta_{\mathrm{R},3}$	$\beta_{\mathrm{R},K}$	$\beta_{\mathrm{B},1}$	$\beta_{\mathrm{B},2}$	$\beta_{\mathrm{B},3}$	$\beta_{\mathrm{B},K}$
Average	95.6	83.2	93.0	136.4	111.6	103.8	106.4	149.4
Selected across all CVs	23	21	24	44	38	34	36	46

Notes: The first line, "Average," represents the average number of nonzero coefficients per card across the five cross-validation splits. The second line, "Selected across all CVs," indicates the total number of coefficients that were retained by LASSO across all five CV splits.

cluded dummy variables indicating whether the red (or black) player has played K (or a number card) n times consecutively, where $n = 4, 5, ..., 29$ (Table B1 in Appendix B.1).

The penalty parameter λ (> 0) is a hyperparameter of the model that is to be determined before selecting the parameter values by solving the maximization problem (6). Following common practice in the machine learning literature, we determine the optimal hyperparameter λ through cross-validation within the training dataset.[8]

The resulting selected values of λ across five cross-validation (CV) rounds are 0.046, 0.060, 0.060, 0.046, and 0.060 for the red players, and 0.060 for all five rounds for the black players.

1. What we can learn from LASSO

Out of 2,448 variables considered, on average (across five CV splits), 408.2 variables were retained for the red players, and 476.0 variables were retained for the black players. These totals represent the sum of nonzero coefficients across the four cards.

Because different rounds of cross-validation involve different selected values of λ and distinct training datasets, the variables chosen by LASSO differ across these five rounds. Given the large number of selected variables, one practical way to interpret LASSO results is to concentrate on variables commonly selected across all five rounds of cross-validation.

The number of variables commonly selected across all CV rounds is 110 for the red player and 154 for the black player; further details are provided in Table 4.

A full list of these commonly selected variables for the red player (card 1) appears in Table 5, and for the red player (card K) in Table 6. Additional related tables are included in Appendix B.3. Parameter values shown in these tables are estimated from the entire dataset.

Table 5—Commonly Selected Variables (Red player, Card 1)

	$\beta_{\mathrm{R},1}$	
	Value	Count
Period dummy (t=6)	0.117	2589
Period dummy (t=27)	−0.064	2589
R played 1 at t-1	−0.435	15988
Action profile at t-1 was (1,3)	−0.077	3206
Action profile at t-1 was (K,1)	−0.044	5386
R played 1 at t-2	−0.116	15980
B played 1 at t-2	−0.083	15792
R played 3 at t-3	0.042	13438
Action profile at t-4 was (3,3)	0.055	2928
R played 1 at t-1, 1 at t-2	0.312	2984
R played 1 at t-1, 2 at t-2	−0.026	3810
R played 1 at t-1, 1 at t-2, 1 at t-3	0.635	720
R played 1 at t-1, K at t-2, K at t-3	0.081	1801
R played 3 at t-1, K at t-2, 2 at t-3	0.103	1156
R played K at t-1, 1 at t-2, 2 at t-3	−0.095	1505
R played K at t-1, 2 at t-2, 3 at t-3	0.120	1177
R played K at t-1, 3 at t-2, 2 at t-3	0.128	1285
R played a number at t-1, a number at t-2, K at t-3, K at t-4	0.065	3171
Winner at t-3 was R	0.054	28155
Winner at t-4 was R	0.021	28093
R lost by playing 1 at t-1, lost by playing 1 at t-2	0.167	953
B lost by playing 1 at t-3	−0.054	6454
B consecutively played numbers in the last 5 periods	0.016	6735

Notes: The 'Value' column presents the estimated coefficients based on the entire dataset. If LASSO with the entire dataset excludes any variable above during the estimation, the corresponding 'Value' entry is left blank. The 'Count' column specifies the number of histories in which each dummy variable is assigned a value of 1.

C. Decision Tree and LightGBM

A *decision tree* partitions data (the pairs of current action of the red or black player and history of play) recursively according to a tree structure, and it returns predictions, one for each subset of data finally obtained by the tree. Figure IV.C shows a tree of depth two. The depth of the tree is defined as the maximum length of paths from the root to the terminal nodes of the tree.

At each node of a decision tree, the question of whether a certain "feature" of the history of play is satisfied or not is asked. The set of candidate features we consider is the set of 2,448 dummy variables that we used for the LASSO model. The depth of a tree, which is a hyperparameter, is chosen by the cross-validation in the training data.[9]

A decision tree works as follows (see Figure IV.C). The feature at node 1 partitions the data into subsets represented by nodes 2 and 3, in each of which the decision tree gives a single probabilistic prediction about the current action given the history of play. The prediction is given by the minimization of the KL diver-

[8]Each set of training data was split into 4 subsets for cross-validation, and the 10 candidate values for λ were chosen from between 10^{-1} to 10^{-2} with equal spacing in the logarithmic scale.

[9]The training data is partitioned into 4 subsets, and candidate depth of 1 to 12 considered. The depth of our model is 5 for the red player and 7 for the black player.

Table 6—Commonly Selected Variables (Red player, Card K)

	$\beta_{\mathrm{R},K}$	
	Value	Count
Period dummy (t=6)	−0.070	2589
Period dummy (t=17)	0.031	2589
Period dummy (t=23)	0.063	2589
Period dummy (t=30)	0.200	2589
Action profile at t-1 was (1,3)	——	3206
B played 2 at t-2	−0.016	13645
B played K at t-2	0.137	24356
Action profile at t-2 was (1,3)	0.088	3195
R played K at t-3	0.074	23380
B played K at t-3	0.197	24427
R played K at t-4	0.135	23435
B played 2 at t-4	−0.028	13619
B played K at t-4	0.191	24480
Action profile at t-4 was (2,2)	−0.097	3115
3-period K-history is ((K,N),(K,N),(K,N))	0.157	686
3-period K-history is ((N,K),(K,N),(N,K))	0.245	772
3-period K-history is ((N,N),(K,N),(N,K))	−0.116	1467
4-period K-history is ((N,K),(N,N),(N,N),(N,N))	0.152	1018
R played K at t-1, a number at t-2, a number at t-3, K at t-4	−0.103	3982
R played a number at t-1, K at t-2, a number at t-3, a number at t-4	−0.087	7366
B played K at t-1, 1 at t-2, 1 at t-3	0.112	978
B played K at t-1, a number at t-2, K at t-3, a number at t-4	−0.099	3738
B played a number at t-1, K at t-2, a number at t-3, a number at t-4	−0.074	7063
Winners at t-1, t-2, t-3 are B, B, B	0.036	13157
Winners at t-1, t-2, t-3 are R, R, R	0.025	4903
Winners at t-1, t-2, t-3, t-4 are B, B, R, R	−0.050	3930
R won by playing K at t-1, won by playing a number at t-2	−0.089	2790
B won by playing a number at t-1, lost by playing K at t-2	−0.105	3092
K-profile and winner were (N,N) and B at t-1	−0.035	10109
R consecutively played K in the last 2 periods	0.089	7792
R consecutively played K in the last 3 periods	0.246	2944
R consecutively played K in the last 6 periods	0.519	392
R consecutively played K in the last 10 periods	1.315	157
R consecutively played numbers in the last 2 periods	0.100	28347
R consecutively played numbers in the last 3 periods	0.009	17752
R consecutively played numbers in the last 4 periods	0.099	11062
R consecutively played numbers in the last 6 periods	−0.073	4359
R consecutively played numbers in the last 7 periods	−0.080	2879
R consecutively played numbers in the last 8 periods	−0.060	1967
R consecutively played numbers in the last 10 periods	−0.192	996
B consecutively played K in the last 3 periods	0.140	3287
B consecutively played numbers in the last 4 periods	0.064	10619
B consecutively played numbers in the last 6 periods	−0.084	4447
B consecutively played numbers in the last 8 periods	−0.183	2175

Notes: The ‘Value’ column presents the estimated coefficients based on the entire dataset. If LASSO with the entire dataset excludes any variable above during estimation, the corresponding ‘Value’ entry is left blank. The ‘Count’ column specifies the number of histories in which each dummy variable is assigned a value of 1.

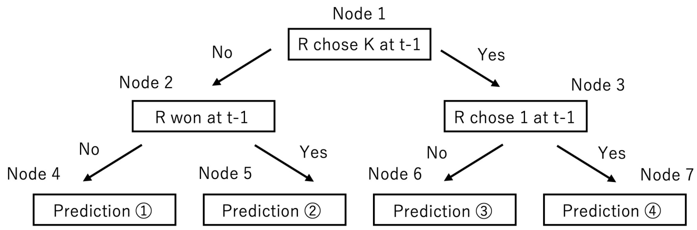


FIGURE 4. AN EXAMPLE OF A DECISION TREE (OF DEPTH 2)

Note: The prediction at each terminal node (leaf) is derived from the frequency of actions in the training data classified to that node.

gence, and it is equal to the empirical frequencies of the current actions in the subset. The feature at node 1 is chosen in such a way that the total loss (prediction errors) at nodes 2 and 3, given the KL divergence, is minimized. Similarly, the feature at node 2 is chosen to minimize the total prediction errors at the succeeding nodes 4 and 5, and so on. The prediction of the model is given by the prediction in the terminal nodes (nodes 4 to 7). This amounts to an automatic selection of the covariates of our general model (1) in the form of the *products* of dummy variables. For example, the final node 4 in Figure IV.C corresponds to a covariate that is equal to (the dummy for the previous action of the red player being K)×(the dummy for the last winner being the red player).

The decision tree literature measures the contribution of each feature used in the tree according to *feature importance*, which we report in Tables 7. Formally, the feature importance of feature x is defined as

$$I(x) = \frac{\sum_{m \in M(x)} \Delta L_m}{\sum_m \Delta L_m},$$

where m is a node, $M(x)$ is the set of nodes associated with feature x (note that the yes/no question about x may be asked at multiple nodes) and

$$\Delta L_m = L_m - L_{(m,\text{Yes})} - L_{(m,\text{No})}$$

is the reduction of total loss (prediction error) at node m. Here, L_m denotes the total loss given by the KL divergence at node m, and $L_{(m,\text{Yes})}$ is the total loss at the succeeding node to m where feature x is true ($L_{(m,\text{No})}$ is similarly defined). Intuitively, the feature importance of x indicates how much the use of feature x improves the prediction of the model. Looking at Table 7, we can see that for both players, the choice probability is affected when the player did not take the action in the past two periods.

We also employed LightGBM, which is an improved version of the decision tree. We considered this model because it has been successful in a wide range of

Table 7—Average feature importance and number of appearances in five CVs

(a) Red players

Variable	Feature Importance	Count
R did not play 1 in the last two periods	0.144	5
R did not play 2 in the last two periods	0.131	5
R did not play 3 in the last two periods	0.095	5
K profile at t-8 is (N,N)	0.087	5
R played K consecutively in the last eight periods	0.068	3
R played K consecutively in the last six periods	0.041	2
K profile at t-3 is (N,N)	0.033	4
R played 2 at t-1 and K at t-2	0.028	3
B played K consecutively in the last three periods	0.025	4
R played 1 consecutively in the last three periods	0.024	3
R played 3 at t-1	0.017	3
B played K at t-3	0.014	2
4-period history of K profiles is ((N,N),(N,N),(N,N),(N,N))	0.013	3
B played K at t-4	0.011	3
B played a number at t-3	0.011	3
R played 2 consecutively in the last three periods	0.010	2
R played 1 at t-1 and K at t-2	0.010	3

(b) Black players

Variable	Feature Importance	Count
B did not play 1 in the last two periods	0.191	5
B did not play 3 in the last two periods	0.090	5
B did not play 2 in the last two periods	0.086	5
B played numbers consecutively in the last eight periods	0.060	5
B played K consecutively in the last four periods	0.048	4
B played numbers consecutively in the last thirteen periods	0.036	5
B won by playing K at t-1 and by playing a number at t-2	0.035	3
K profile at t-1 is (N,K)	0.033	4
B played numbers consecutively in the last nine periods	0.020	3
B played 3 at t-1 and 1 at t-2	0.020	3
B played K consecutively in the last seven periods	0.019	5
B played K consecutively in the last three periods	0.016	1
B won by playing K at t-1 and by playing 2 at t-2	0.013	2
B played 3 at t-1 and K at t-2	0.012	1
B played a number at t-1 and K at t-2	0.011	2
B played 1 at t-1 and t-2	0.010	3
K profile at t-4 is (N,N)	0.010	4

Notes: Average feature importance and the number of appearances (duplicates within a single tree do not count) in five trees for (a) the red player and (b) the black player. Note that we report only variables whose feature importance is greater than 0.01.

applications. Its performance in our study, however, turned out not to be among the best, and therefore, we skip the formal definition of this model, which can be found in Ke et al. (2017).

D. Deep Learning Models

We consider two types of deep learning algorithms: the basic Deep Neural Network (DNN) and Long-Short Term Memory (LSTM). The DNN model is the simplest form of deep learning. We created a five-layer neural network composed of one input layer, one output layer, and three hidden layers, as described in Figure 5. The input to the machine is the four-period history of (the action profile and the payoffs to the player), and the output is the vector of choice probabilities of the action in the current period.[10] That is, each input is a vector of $(16+1)\times 4 = 68$ dummies across four periods, and each output is a four-dimensional probability vector.[11] The number of cells in the three hidden layers is a hyperparameter that is determined by cross-validation in the training data, where the number of cells in each hidden layer is optimally selected from the candidate set $\{10, 30, 50\}$.[12] Each cell in the hidden and output layers is densely connected in the sense that it is connected to all cells in the previous layer.

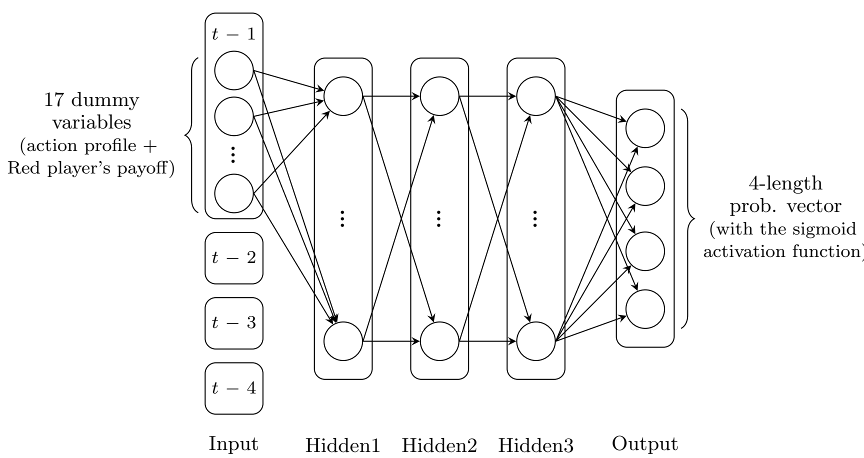


FIGURE 5. AN ILLUSTRATION OF A DNN MODEL

DNN recursively determines the value of each cell in the following way. Let j be a cell in the hidden or output layer, and let $i = 1, ..., I$ be the cells in the previous

[10]Even if the payoff to the focal player can be deduced from the action profile, including the payoff data helps to reduce the prediction error of the model.

[11]For each period, we used 16-dimensional dummy variables of action profiles and 1-dimensional payoff of the red player as input data.

[12]The selected set of parameters is $(30, 50, 50), (50, 30, 50), (50, 30, 30), (50, 30, 30), (50, 50, 10)$ for the red players, and $(50, 30, 30), (50, 50, 50), (50, 30, 30), (50, 30, 30), (50, 30, 30)$ for the black players.

layer. The branch from i to j is associated with parameter (or weight) w_{ij}, and given the values x_i of previous cells $i = 1, ..., I$, the input to cell j is determined as $\sum_i w_{ij} x_i$ and the value of j, denoted x_j, is determined by activation function f as $x_j = f(\sum_i w_{ij} x_i)$. We adopt the ReLU activation function $f(x) = \max\{0, x\}$ for all cells except those in the output layer. Cells in the output layer use the softmax activation function (multinomial logit). The parameters w_{ij} are selected to minimize the loss defined by the KL divergence. Note that the machine has a large number of parameters w_{ij}, and the meaning of the value of each parameter is not immediately clear. This is why the deep learning model is often called a "black box," and this fact will play an important role in our analysis presented below.

We used the PyTorch implementation of a standard DNN. To avoid overfitting, we used early stopping and dropouts when training the model. We trained our model with the Adam optimizer.

LSTM is a Markovian version of the deep learning model, where state variables are updated in a Markovian way, and the prediction for each period is given by the current values of the state variables. The neural network architecture is used to implement the state transition and the derivation of predictions from the state variables. LSTM has been successfully used in a number of applications, such as Google and Facebook machine translation systems, Google's voice recognition application, and autocorrect on iOS.

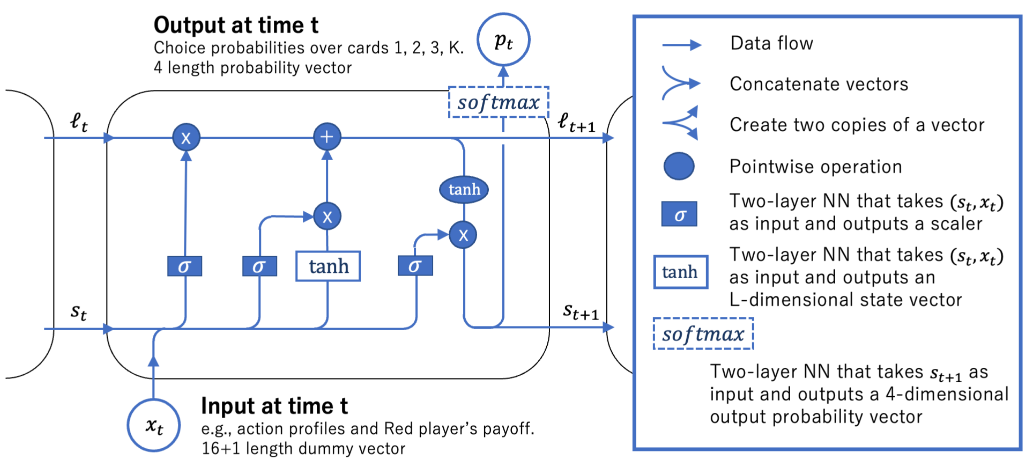


FIGURE 6. AN UNFOLDED ILLUSTRATION OF THE LSTM MODEL.

Note: An unfolded illustration of the structure of our LSTM model architecture based on Gers, Schmidhuber and Cummins (2000). This figure is drawn by the authors based on Olah (2015).

We use the PyTorch implementation of LSTM, which is a standard network introduced in Gers, Schmidhuber and Cummins (2000). Our LSTM model is illustrated in Figure 6. This model includes two state vectors, s_t and ℓ_t, both of which are numerical vectors of dimension L. Here, L is a hyperparameter of the model and is determined by the cross-validation in the training data. As a result, we adopted $L = 20$ to 80 (50 to 70 for 8/10 splits) for the five training data sets from the candidate dimensions $\{10, 20, \ldots, 80\}$.

The LSTM module for period t (encompassed by a rounded rectangle in the figure) takes two state vectors and a new data vector (the action profile in and the win/loss of the player in the previous round) as inputs and outputs new state variables and choice probabilities for the action in period t.

We trained our model using the Adam optimizer. To avoid overfitting, we used early stopping and dropouts when training the model.

V. Performance Comparison

We compared the model performance using leave-one-out cross-validation as discussed in Section IV. We first randomly divided the entire 2589 pairs in our dataset into five subgroups $D_1, \ldots, D_5$ of nearly equal size. Four of them concern 2018 pairs, while the remaining one concerns 2017 pairs. Then, for each $k = 1, \ldots, 5$, we trained (estimated) each model using its training data $\cup_{\ell \neq k} D_\ell$ and computed the out-of-sample prediction errors (the KL loss) using its test data D_k.

In particular, we computed the error rate of each model trained in each round of cross-validation as follows. Since the predictions of our machine learning models depend on the history of play in the last four periods, they provide predictions for periods $t = 5, 6, \ldots, 30$. Therefore, we first compute, in each round of cross-validation, the average error (per observation) of each model as (the sum of prediction errors given by the KL loss (5) over all observations) divided by (the number of observaions $|D_k| \times 26$). Next, we take the average of those numbers across the five rounds of cross-validation. These are the error rates of the models we report in what follows.

The error rate given by the KL divergence for our best model is, as we will discuss shortly, 1.331, but this number is hard to interpret. To address this issue, we introduce the notion of the *strategic error rate*, which is easy to interpret and defined by the following loss function

$$l(P, q) = 1 - \pi_{-i}(\mathrm{BR}_{-i}(P), q), \tag{7}$$

where π_{-i} is the payoff function of the opponent of player i, which takes on value 1 for a win (and is otherwise equal to zero), BR_{-i} denotes the best reply of the opponent of player i, and finally, P and q are the prediction of the model and the degenerated probability distribution representing the action choice of player i. The strategic error rate, the average loss (7) across all observations, provides the answer to the following question. Consider a model of, for example, the red player. Suppose we randomly select one pair of players from the data and also randomly select a round denoted by t. Given the history of play of the pair up to round t, the model gives a probabilistic prediction of the red player's action in round t. If the black player utilizes that prediction of the model, takes the (myopic) best reply, and faces the actual choice of action of the red player, on average, how often does the black player lose? For example, if the strategic error rate of a model is 0.35, the opponent loses 35% of the time if she utilizes the model of the player and takes the best reply in any single period.[13]

A. Results

Table 8—Prediction Performance (Conventional and Machine Learning Models)

	Red Players				Black Players			
	#Params	KL	SER	RC	#Params	KL	SER	RC
Conventional Models								
Constant (Baseline)	3	1.360	0.350	0.000	3	1.354	0.562	0.000
2-period Nonparametric	768	1.354	0.341	0.215	768	1.347	0.539	0.297
EWA	8	1.352	0.332	0.276	8	1.349	0.559	0.195
Serial Correlation								
Order 1	5	1.354	0.350	0.204	5	1.349	0.541	0.199
Order 4	11	1.351	0.343	0.293	11	1.346	0.536	0.334
Machine Learning Models								
LASSO	408.2	1.344	0.324	0.562	476.0	1.339	0.526	0.589
Decision Tree	40.4	1.352	0.343	0.273	41.6	1.348	0.536	0.248
LightGBM	30.0	1.344	0.325	0.547	30.0	1.340	0.527	0.556
DNN	6208.0	1.347	0.332	0.455	6428.0	1.343	0.534	0.432
LSTM	5524.0	1.331	0.308	1.000	11204.0	1.329	0.507	1.000
Human			0.419				0.581	

Notes: Average prediction performance of the selected models measured in the test data. KL, SER, and RC denote Kullback-Leibler divergence, strategic error rate, and relative completeness. Each performance score is the average of five train-test CV splits. The average training (test) data size is 2071.2 (517.8) observations. Hyperparameters of the machine learning models are determined by cross-validation within training data. The number of model parameters (#Params) is the average of the five models for each split. The performance table of the other models is shown in Appendix C.1. The scores of the other performance measures (L1 and L2 loss measured on test data, and AIC and BIC on the training data) are listed in Appendix C.2.

The performance comparison results are presented in Table 8, which reports the KL divergence and the strategic error rates of our leading models on the test

[13]We stress that the strategic error rate concerns the usefulness of the estimated model in any given *single* period. In principle, if the back player in the previous story makes the best reply to the estimated model of the red player in one period, this has two effects: (i) the loss rate in the current period changes, and (ii) the future behavior of the red player changes, which affects the future loss rates of the black player. Our notion of strategic error rate only measures the former effect.

data, averaged across five CV splits. Note that all our models were trained to minimize the KL divergence on the training data.

Table 8 shows that, among the four machine learning model — LASSO, decision tree, LightGBM, deep learning models DNN and LSTM — LSTM has outstanding performance, followed by LASSO. LSTM takes on a flexible functional form of the deep learning model, and its Markovian structure is suitable for processing sequential data.[14]

Given this, we measure the relative predictive powers of our models in terms of the error rate reduction from a naïve benchmark in the form of an i.i.d. mixture model (constant baseline) to the best model LSTM.[15] This leads to the notion of *relative completeness* (RC) defined as

$$\text{(RC of a model)} = \frac{\text{(KL of the model)} - \text{(KL of Constant)}}{\text{(KL of LSTM)} - \text{(KL of Constant)}},$$

where KL (KL divergence) is the average of the five CV (cross-validation) splits. This is a version of the *completeness* measure introduced by Fudenberg et al. (2021). If we replace LSTM (the best model we have obtained) with the true best model (the best one among all specifications of the model Equation (1)) in the above definition of RC, we obtain their *completeness* measure.[16]

We observe that EWA, a leading learning model in behavioral economics, has a relative completeness of 0.286 (0.192) for the red (black) player. This means that for the red (black) player, EWA achieves only 28.6% (19.2%) of the predictive power of the best machine learning model LSTM (in terms of the error rate reduction from the naive benchmark i.i.d. mixture model). The other class of behavioral models, the serial correlation, exhibit performance similar to that of EWA (relative completeness of 0.210-0.302 (0.202-0.340) for the red (black) player). This shows that the traditional behavioral models only have 20-30% of the predictive power of the best machine learning model. They are missing out on some regularities of human behavior, and there is room to improve those models. The notion of RC, or the relative error rate reduction, provides a useful benchmark to intuitively understand how much the traditional models can be improved.

Let us now turn to the strategic error rates. In the mixed strategy equilibrium, the loss rate for the black player is 40%, and this is true (by the property of

[14]However, traditional econometric criteria of model selection, such as AIC and BIC in the training sample, do not select LSTM due to its large parameter size. Machine learning models DNN and LSTM are known to be "over-parametrized" in the sense that different parameter configurations can provide the same or similar functional form. Hence, these traditional information criteria that penalize the number of parameters are known to be inappropriate in the machine learning literature. We list the average AIC and BIC of the representative models in Appendix C.2.

[15]The naïve benchmark model of a player assumes that the player's action is i.i.d. over time according to the empirical frequencies of actions in the training data.

[16]To estimate the true best model among all specifications is impossible in our setting. Since there are a total of 16^t types of t-period histories, a fully non-parametric model should specify the choice probability for $\sum_{t=0}^{29} 16^t$ kinds of histories.

completely mixed strategy equilibrium) no matter which action the black player chooses. If, in our dataset set, the black player utilizes EWA or the serial correlation model of order 4 to predict the red player's choice and takes the best reply action, in contrast, the loss rate can be reduced to 33.1% or 34.3%. The loss rate of the black player can be further reduced to 30.8% if she/he utilizes the best machine learning model, LSTM.[17] This shows that the prediction error of the best machine learning model (LSTM) is significantly lower than those of the traditional models in a strategic sense. The statistical significance of the difference in prediction errors will be discussed in Subsection VI.C.

Table 8 also shows that the non-parametric model that depends on two-period histories of action profiles (2-period Nonparametric) fares worse than EWA and the serial correlation model and much worse than the best machine learning model, LSTM. We have a sufficiently large data set for two-period non-parametric estimation, where each 2-period history occurred at least 66 times (the average occurrence is 291.9). This provides fairly reliable evidence that the subjects' memory is longer than two periods.

B. The Role of Big Data

In this section, we show that the merits of the machine learning models can be seen only when the data size for parameter fitting is an order of magnitude larger than that of the typical lab dataset.

To this end, we show how the out-of-sample prediction errors of various models change if we conduct parameter fitting in artificially reduced data sets (while keeping the data size for performance comparison fixed).[18] The outcomes are shown in Figure 7 (ME1, ME2, and ME2* in the figure are the modified models that we introduce later). The figure shows that the performance of the traditional models, Constant, EWA, and Serial Correlation, do not improve if we increase the training data size beyond 400 pairs. In contrast, the error rates of the machine learning models, LASSO and LSTM, keep on decreasing as the training data size increases, and at our dataset size of 2,000+ pairs, the difference in out-of-sample prediction errors between the best machine learning model, LSTM, and the conventional models is quite evident. In contrast, if we have 400 pairs of subjects, our best machine learning model, LSTM, is only marginally better than the traditional models, and if the number of pairs is less than 200, LSTM fares worse than the traditional models.

The reason why we need a large data set to see the merit of the machine learning model is that the machine learning model has a large number of parameters, as

[17] One may note that, for the red player, the serial correlation of order 4 model has a larger strategic error rate than EWA, but its KL divergence is smaller than that of EWA. This happens for the following reasons. Given a history of the play, consider the empirical frequencies of actions (call it (a)) and prediction of action distribution by EWA (b) and by the serial correlation (c). It can happen that, even though (a) and (c) are close to each other (low KL), (a) and (b) may have the same best reply, but (c) does not. Then, (b) fares better than (c) in terms of the strategic error rate.

[18] Peterson et al. (2021) conducted a similar analysis.

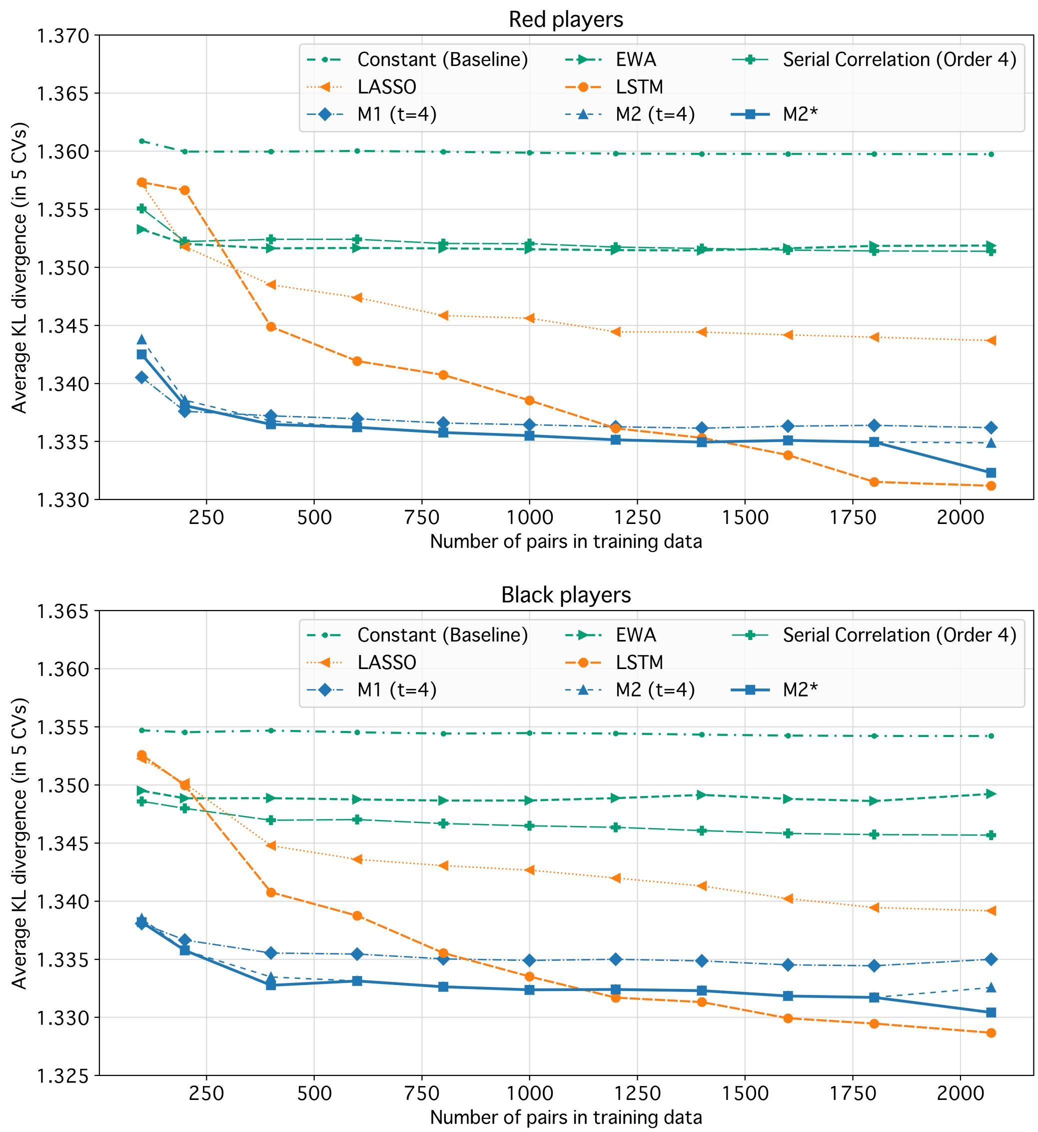


FIGURE 7. AVERAGE KL DIVERGENCES WITH SMALLER TRAINING SAMPLES

Note: Average KL divergences of machine learning models compared to traditional models as the size of the training dataset is artificially reduced. In most laboratory experiments, the number of observations is typically under 6,000 (see Figure 2), corresponding to around 100 to 200 pairs in the plots. Improved models, ME1, ME2, and ME2* will be explained in Section VI.B. They demonstrate performance that is very close to the best machine learning model, LSTM, when using the full training sample. Furthermore, unlike LSTM, their performance remains good even with smaller sample sizes.

shown in Table 8. As we noted in the introduction, however, a large number of parameters of a model does not necessarily guarantee its high model performance. If the subjects' behavior were independent of the past history of the play, for example, the model with only three parameters (constant terms for 1, 2, and 3), which predicts player behavior according to the empirical frequency of actions in the training data, would have much lower out-of-sample prediction errors than any other model with a larger number of parameters. In other words, models with too many parameters should fare worse because they "overfit" the data used for parameter fitting. Therefore, our finding that the best machine learning model, which has a large number of parameters, fares much better than the traditional model in our big dataset is by no means a trivial observation that can be expected *a priori*. It shows that our best machine learning model has successfully captured some regularities of human behavior that the traditional models have failed to detect, and this can only be seen when the dataset size is much larger than usual.

VI. Opening the Black Box of the Machine Learning Models to Improve the Traditional Models

A. Procedure

Our next task is to improve the traditional behavioral models by incorporating what is captured by the machine learning models. The challenge is, however, that a machine learning model is usually a black box that is difficult to interpret. This is especially true for our best-performing machine learning model, LSTM. The parameters of this model are the weights attached to more than 5,000 branches in its network structure, whose meaning is hard to interpret. To address this issue, we adopted the following procedure to open the black box.[19] First, we examined more interpretable machine learning models we estimated, the decision tree and LASSO models. In particular, we examined the feature importance of the decision tree model and the commonly selected right-hand side variables in the five rounds of cross-validation of the LASSO model. Second, we modified the conventional behavioral models by incorporating the insights gained from those observations. We obtained new behavioral models, which we call Mixing EWA models, that capture how humans learn to play mixed strategies. Third, we examined how much out-of-sample predictive power the improved models have compared to the best machine learning model. We indeed find that there is no statistically significant difference in the predictive power of our best modified behavioral model and the best machine learning model. Fourth, we conducted two statistical tests to see if our modified models actually capture what is encoded in the black box of our best machine learning model ("decoding verification"). Finally, we point out a potential concern about our procedure, and to address

[19]Opening the black box of machine learning is a hot research topic in AI under the rubric of XAI (explainable AI). Decoding our machine learning models by the XAI techniques is an intriguing agenda for future research.

this issue, we conduct a double-checking of the external validity of our modified model.

B. Improved Models

The feature importance of our decision tree model shows that an important factor in determining current action choice is whether a certain action was consecutively not played in the last two periods (Table 7). The lists of the variables commonly selected in the five rounds of cross-validation of the LASSO model are shown in Table 5, 6, and Tables B2-B7 in Appendix B.3. It suggests that the choice of a card is affected by (1) whether the card was chosen in the previous period and (2) whether the card was consecutively chosen (or not chosen) in the last four periods. This points to the importance of a very specific form of serial correlation based on the *repeated* choice or *repeated* avoidance of a certain set of actions. To capture this, let us introduce the following variables. For any subset $B \subseteq \{K, 1, 2, 3\}$, let $R_i^t(B, \tau)$ be a vector of indicator functions that show player i repeatedly chose (or avoided) actions in B:

$$R_i^t(B, \tau) = \begin{pmatrix} 1\Big\{a_i^{t-s} \in B, s = 1, ..., \tau\Big\} \\ \\ 1\Big\{a_i^{t-s} \notin B, s = 1, ..., \tau\Big\} \end{pmatrix}.$$

For a singleton set $B = \{a\}$, we write $R_i^t(B, \tau)$ as $R_i^t(a, \tau)$.

Using those variables, we obtain what we call *Mixing EWA* models that try to improve upon the conventional behavioral models. The first version, Mixing EWA 1, or ME1 for short, incorporates a specific form of serial correlation into the EWA model and is formulated as follows.

$$P_i^a(t) \propto \exp\Bigg\{\lambda A_i^a(t-1) + \sum_{\tau=1}^{4} \alpha_i^{a\tau} \cdot R_i^t(a, \tau)\Bigg\},$$

where $A_i^a(t-1)$ is the EWA attractor defined by Equation (4). We assume that the coefficient vectors in the above formula are common for all number cards ($\alpha_i^{1\tau} = \alpha_i^{2\tau} = \alpha_i^{3\tau}$) but can differ from those for K. This is because this parameter restriction has provided the best predictive performance.[20] Analogous parameter restrictions apply for the second modified model that will be introduced in what follows. Table 9 shows the out-of-sample predictive power of ME1. The relative completeness of ME1 is 83% for the red player and 75% for the black player, which are substantial improvements over the conventional models.

The relative completeness of ME1 shows that this improved model is still missing out on some important empirical regularities that are captured by our best ma-

[20] We compare the performance of the restricted and unrestricted models in Appendix C.1

Table 9—Prediction Performance (Improved Models)

	Red Players				Black Players			
	#Params	KL	SER	RC	#Params	KL	SER	RC
Conventional Models								
Constant (Baseline)	3	1.360	0.350	0.000	3	1.354	0.562	0.000
EWA	8	1.352	0.332	0.276	8	1.349	0.559	0.195
Serial Correlation (Order 4)	11	1.351	0.343	0.293	11	1.346	0.536	0.334
Improved Models								
ME1 (t=4)	24	1.336	0.314	0.825	24	1.335	0.521	0.753
ME2 (t=4)	38	1.335	0.311	0.870	38	1.333	0.517	0.848
ME2*	126	1.332	0.311	0.961	110	1.330	0.516	0.932
Best ML Model								
LSTM	5524.0	1.331	0.308	1.000	11204.0	1.329	0.507	1.000

Notes: Average prediction performance of the selected models measured in the test data. KL, SER, and RC are the abbreviations for Kullback-Leibler divergence, strategic error rate, and relative completeness. Each performance score is the average of five train-test CV splits. The average training (test) data size is 2071.2 (517.8) observations. The number of model parameters (#Params) is the average of five models. The scores of the other performance measures (L1 and L2 loss measured in the test data, and AIC and BIC measured in the training data) are listed in Appendix C.2. ME2* is the ME2 model whose history length t minimizes the KL divergence among all $t = 1, 2, \ldots, 29$. Here, the best t is 15 for the red player and 13 for the black player.

chine learning model. If the particular form of serial correlation captured by ME1 is important, a subject may pay attention to the fact that the opponent's behavior is so affected. Moreover, when player i considers choosing card a, what is relevant is whether the opponent chooses the particular cards against which player i's card a wins. To capture this idea, let W_i^a be the set of opponent cards for which a player i wins by choosing card a. Formally, $W_i^a = \{a' \in \{1, 2, 3, K\} \mid \pi_i(a, a') = 1\}$, where $\pi_i(a, a')$ denotes player i's payoff when she plays a and her opponent plays a'. Our second modified model, Mixing EWA2, or ME2 for short, is defined as follows: for each $a = K, 1, 2, 3$,

$$P_i^a(t) \propto \exp\left\{\lambda A_i^a(t-1) + \sum_{\tau=1}^{T} \alpha_i^{a\tau} R_i^t(a, \tau) + \sum_{\tau=1}^{T} \gamma_i^{a\tau} R_{-i}^t(W_i^a, \tau)\right\}, \tag{8}$$

where $T = 4$. To avoid multicollinearity, we set $\gamma_i^{a1} = (\gamma_{i1}^{a1}, 0)$.[21]

Table 9 shows ME2 captures 87% (for the red player) and 85% (for the black player) of the predictive power of our best machine learning model LSTM. We then varied T in Formula (8) to see which value of T minimizes the out-of-sample

[21] There are multicolinearites in the elements in vector $R_i^t(a, 1)$ and vector $R_{-i}^t(W_i^a, 1)$. This is because $(1,1)R_i^t(a, 1) = 1\left\{a_i^{t-1} = a\right\} + 1\left\{a_i^{t-1} \neq a\right\} = 1$, and similarly $(1,1)R_{-i}^t(W_i^a, 1) = 1$. Hence the second element of the vector $R_{-i}^t(W_i^a, 1)$, namely $1\left\{a_{-i}^{t-1} \notin W_i^a\right\}$, is to be omitted.

prediction error. We show the KL divergence of each ME2(T) moels in Table D12 in Figure D3. We found that the optimal value is $T = 15$ for the red player and $T = 13$ for the black player. Let us call those models ME2*. Remarkably, this model captures almost all predictive power (96% for the red player and 93% for the black player) of our best machine learning model.

By incorporating the insights gained from our machine learning models, we have obtained improved models that are applicable to learning behavior not just in the O'Neill's game but in general games with a unique mixed strategy equilibrium. The Mixing EWA models, ME1 and ME2, modify the leading model of learning, EWA, in the following ways. EWA captures two important factors that guide a player trying to learn how to play a game. One is which strategy fared well in the past history of play, and the other is the prediction of the opponent's strategy based on the past history. However, a player's desire to make one's own behavior unpredictable in a game with a unique mixed strategy equilibrium can only be mechanically captured by the i.i.d. logit noise in the model, despite the fact that humans are not good at creating i.i.d. sequences. Our Mixing EWA makes this aspect of EWA more realistic by incorporating a player's naive attempt to make oneself unpredictable by choosing (or avoiding) the same strategy repeatedly. ME2 is more sophisticated than ME1, and it allows a player to pay attention to the opponent's naive attempt to be unpredictable. While ME1 applies to all games with a unique mixed strategy Nash equilibrium, ME2 can be applied to a subset of such games, zero-sum two-person games with binary outcomes, either "player 1 wins" or "player 2 wins." We hope these models prove useful in explaining how human players learn to play mixed strategies.

Figure 7, which we have seen previously, also shows how the out-of-sample prediction errors of the Mixing EWA models change as the training data size for parameter estimation varies. ME2* in the figure refers to the ME2 models with the optimal value of T for each training data size. It shows that the Mixing EWA models perform well over the entire range of data sizes, including the range of data sizes typically found in lab experiments.

C. *How to Conduct Statistical Tests on the Cross-validation Outcomes*

The performance comparison of our models in Tables 8 and 9 is based on cross-validation (CV). Recall that it works as follows. First, we randomly split the whole dataset into K subsets of an (almost) equal size, calling all of them $D_1, \ldots, D_K$. Second, we set aside D_k as test data, conduct parameter fitting of our models on the remaining training data $\bigcup_{h \neq k} D_h$, and then measure their out-of-sample prediction errors in the test data D_k. This process is repeated for $k = 1, \ldots, K$. Finally, models are compared according to the *average* out-of-sample prediction errors in the K rounds of CV.

Are the differences in the prediction errors (KL) of our models reported in Tables 8 and 9 statistically significant? This is not a trivial question because the reported errors are the *average* taken over the $K(= 5)$ rounds of CV, where the estimated parameters of the models are *different* across the K rounds.

To address this issue, we suggest two kinds of tests, an *overall test* and a *round-wise test*, which serve somewhat different purposes. The overall test compares the reported prediction errors in CV, the average errors over the K rounds. This test was developed by Austern and Zhou (2024) and introduced into economic analysis by Fudenberg, Gao and Liang (2023).[22] It concerns the expected out-of-sample prediction errors of models when their parameter fitting is conducted in a given size of data. Suppose we have data with size N and conduct K rounds of CV. Then, the training data size for parameter fitting is $(K-1)N/K$. Austern and Zhou is concerned with the expected out-of-sample prediction errors of models when parameter fitting is conducted in a data set whose size is $(K-1)N/K$. Consider a model with parameter space Θ, denoted by $\mathcal{F}_\Theta$. The expected out-of-sample error when this model is optimally estimated in the data of size $(K-1)N/K$ is denoted by $e_{\mathcal{F}_\Theta, \frac{K-1}{K}N}$. The null hypothesis we test is whether the expected error is the same for LSTM and any other given model. Intuitively, it addresses the following kind of question. If we randomly sample a new data set of size $(K-1)N/K$ that comes from the true underlying distribution, and conduct parameter fitting of various models using the data, what are the expected out-of-sample prediction errors of the models? The precise description of the Austern-Zhou test, which we call the overall test, is given in Appendix A, and the results are reported in Table 10. The test in the table examines the statistical significance (in the above-mentioned sense) of the difference in prediction errors (KL) between the best model (LSTM) and each of the other models. The p-values in the table strongly show that there are significant differences in the prediction errors of the best model (LSTM) and the rest.

Let us now discuss the difference between a *model specification* and an *estimated model*. Model specification refers to a functional form, such as EWA, with parameters whose values are to be estimated. In contrast, an estimated model refers to a model with estimated parameters. The overall test is all about model specifications and is suitable for answering the following question. Which model specifications will provide better out-of-sample predictions if we have a new data set for parameter estimation, whose size is equal to that of our current training data? If our purpose is to select the best model specification to make an accurate prediction, the overall test is the one we should use. It also provides a handy way to see the statistical significance of the cross-validation results.[23]

However, our goal in this paper is somewhat different. We want to see if the *estimated* models of machine learning are better than the traditional models. If the answer is yes, we can then try to improve the traditional models using the insights gained from the *estimated models* of machine learning. This observation

[22]We are grateful to Annie Liang for helpful suggestions and discussions.

[23]The validity of the overall test is currently established under a certain set of rather restrictive conditions (see Appendix A), which may not be satisfied by all of our models. In particular, strict convexity of the loss function of a mode may not be true for the machine learning models. The roundwise test, in contrast, does not have such a potential concern.

Table 10—Overall test for prediction performance

	Red Players			Black Players		
	KL	Diff	p-value	KL	Diff	p-value
Conventional Models						
Constant (Baseline)	1.360	0.029	1.57×10^{-94}	1.354	0.026	1.01×10^{-90}
EWA	1.352	0.021	2.04×10^{-55}	1.349	0.021	1.25×10^{-70}
Serial Correlation						
Order 1	1.354	0.023	2.33×10^{-68}	1.349	0.020	3.26×10^{-64}
Order 4	1.351	0.020	2.18×10^{-61}	1.346	0.017	1.94×10^{-53}
Machine Learning Models						
LASSO	1.344	0.013	2.00×10^{-31}	1.339	0.010	8.74×10^{-24}
Decision Tree	1.352	0.021	2.69×10^{-65}	1.348	0.019	5.92×10^{-58}
LightGBM	1.344	0.013	2.22×10^{-34}	1.340	0.011	2.58×10^{-28}
DNN	1.347	0.016	3.08×10^{-52}	1.343	0.014	1.92×10^{-49}
LSTM	1.331	0.000	—	1.329	0.000	—
Improved Models						
ME1 ($t = 4$)	1.336	0.005	9.93×10^{-7}	1.335	0.006	3.66×10^{-9}
ME2 ($t = 4$)	1.335	0.004	2.44×10^{-4}	1.333	0.004	2.35×10^{-4}
ME2*	1.332	0.001	2.96×10^{-1}	1.330	0.002	1.62×10^{-1}

Notes: The Diff column indicates the difference between the average KL divergence of the LSTM model and the average KL divergence of each respective model. The p-value column represents the p-value of the Austern-Zhou test discussed in Section VI.C. The null hypothesis states that the expected KL divergence of each model, estimated in the population, is equal to that of the LSTM model trained in the population.

leads us to the *roundwise test* of our CV results. This test asks if the differences in the out-of-sample prediction errors of the estimated models in each round of CV are statistically significant. It is a simple standard test about the null hypothesis that (prediction error of a given estimated model) − (prediction error of the best estimated model LSTM) has mean 0, and the results are reported in Table 11. The roundwise test shows that there are significant differences in prediction errors between the estimated LSTM models and the rest (other than the improved models), which validates our presumption that we can improve the existing models by examining the estimated models of machine learning. In contrast, statistical significance differs across the rounds of CV for the improved models. In some rounds, there are no statistically significant differences in the prediction errors of the estimated LSTM and the estimated improved models.

Table 11—Roundwise test for prediction performance (KL divergence, each CV)

	Red Players					Black Players				
	CV1	CV2	CV3	CV4	CV5	CV1	CV2	CV3	CV4	CV5
Conventional Models										
Constant (Baseline)	1.361^{***}	1.359^{***}	1.362^{***}	1.361^{***}	1.355^{***}	1.359^{***}	1.351^{***}	1.352^{***}	1.354^{***}	1.356^{***}
EWA	1.351^{***}	1.352^{***}	1.350^{***}	1.355^{***}	1.350^{***}	1.353^{***}	1.347^{***}	1.347^{***}	1.348^{***}	1.351^{***}
Serial Correlation										
Order 1	1.358^{***}	1.352^{***}	1.355^{***}	1.355^{***}	1.349^{***}	1.353^{***}	1.345^{***}	1.347^{***}	1.351^{***}	1.349^{***}
Order 4	1.357^{***}	1.350^{***}	1.352^{***}	1.352^{***}	1.346^{***}	1.350^{***}	1.341^{***}	1.343^{***}	1.347^{***}	1.347^{***}
Machine Learning Models										
LASSO	1.347^{***}	1.346^{***}	1.342^{***}	1.344^{***}	1.339^{***}	1.344^{***}	1.336^{***}	1.335^{***}	1.340^{***}	1.341^{***}
Decision Tree	1.356^{***}	1.352^{***}	1.350^{***}	1.354^{***}	1.348^{***}	1.354^{***}	1.347^{***}	1.346^{***}	1.347^{***}	1.345^{***}
LightGBM	1.347^{***}	1.345^{***}	1.343^{***}	1.345^{***}	1.339^{***}	1.344^{***}	1.337^{***}	1.336^{***}	1.342^{***}	1.341^{***}
DNN	1.351^{***}	1.348^{***}	1.346^{***}	1.347^{***}	1.343^{***}	1.347^{***}	1.338^{***}	1.341^{***}	1.344^{***}	1.345^{***}
LSTM	1.339	1.332	1.326	1.333	1.325	1.336	1.326	1.324	1.326	1.332
Improved Models										
ME1 ($t = 4$)	1.339	1.337^{***}	1.329	1.341^{***}	1.335^{***}	1.339	1.333^{***}	1.331^{***}	1.336^{***}	1.336^{*}
ME2 ($t = 4$)	1.338	1.337^{**}	1.328	1.339^{***}	1.333^{***}	1.336	1.331^{***}	1.329^{**}	1.334^{***}	1.334
ME2*	1.335	1.337^{**}	1.324	1.337^{*}	1.329^{**}	1.333	1.330^{**}	1.326	1.331^{**}	1.332

Notes: We compared the KL divergence of each estimated model with that of the LSTM model using paired t-tests for each of the five cross-validation splits. The null hypothesis is that there is no difference in the KL divergence between a given model and the LSTM model, meaning the mean difference is zero. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

D. *Double-checking the External Validities of the Improved Models*

Strictly speaking, the above performance comparison is not entirely "fair" as it can potentially favor the modified models. This is because the modified models are constructed from the better-performing models in the test data, and therefore, the functional forms of the modified models utilize the information in the test data (although parameter fitting is done in the training data). In other words, the improved model may overfit the test data, and its good performance may come from the goodness of fit to the test data, not from the fact that it captures the true regularities of the subjects' behavior.

To address this potential problem, we conducted a performance comparison by means of a new dataset, which we collected after obtaining the data we used in our preceding analysis in this paper. The new dataset comes from April 2021 to March 2023 and has 231 pairs of observations. This is about half of the average number of pairs in the test data used in the cross-validation (517.8).

We conducted a prediction contest once, using all prior data for parameter fitting and the new data for performance comparison. The results are shown in Table 12. KL' and RC' are the KL divergence and the relative completeness obtained from the new data. The table shows that the results are largely unchanged. The relative ranking of models is unchanged except that EWA fares worse than the baseline for the black player.[24] The stars in the table refer to the simple test that we used for our roundwise test because we did not use cross-validation. Our best improved model (ME2*) has, for the red player, an out-of-sample prediction error that is not significantly different from that of the best machine learning model, LSTM, at the 5% level.

E. *Decoding Verification*

Our best modified model, ME2*, has 93-96% of the predictive power of the best machine learning model, LSTM. This could mean that ME2* successfully decoded the black box of LSTM, but there is a possibility that ME2* and LSTM capture different regularities of data and happen to perform similarly.

To examine if ME2* successfully decodes what is captured by our best machine learning model (LSTM), we propose what we call *residual tests*. First, we define the "attractors" of our models. Note that both ME2* and LSTM have the property that the choice probability of card $a = K, 1, 2, 3$ by player i at time t is given by

$$P_i^a(t) = \frac{\exp\left(W_i^a(t-1)\right)}{\sum_{a'=K,1,2,3} \exp\left(W_i^{a'}(t-1)\right)},$$

for some scalar $W_i^a(t-1)$, which we call the "attractor." Since the probability

[24]For the black player, EWA and ME1 have lower performance in the new data, and it may be because the small test data happened to contain histories that do not favor the EWA model.

Table 12—Prediction Performance (Existing Data vs New Data)

	Red Players				Black Players			
	KL	KL’	RC	RC’	KL	KL’	RC	RC’
Conventional Models								
Constant (Baseline)	1.360	1.362***	0.000	0.000	1.354	1.351***	0.000	0.000
EWA	1.352	1.355***	0.276	0.187	1.349	1.353***	0.195	−0.045
Serial Correlation ($t = 4$)	1.351	1.350***	0.293	0.294	1.346	1.338***	0.334	0.380
Improved Models								
ME1 ($t = 4$)	1.336	1.334***	0.825	0.681	1.335	1.334***	0.753	0.510
ME2 ($t = 4$)	1.335	1.328***	0.870	0.810	1.333	1.330***	0.870	0.638
ME2*	1.332	1.326*	0.961	0.868	1.330	1.325***	0.932	0.763
Best ML Model								
LSTM	1.331	1.320	1.000	1.000	1.329	1.317	1.000	1.000

Notes: Average prediction performance scores are measured on the test data. The KL and RC columns are the KL divergence and relative completeness based on the old data (averaged across five CV splits), while KL’ and RC’ are based on the new data. For KL and RC, the average training and test dataset sizes are 2,071.2 and 517.8 observations, respectively; for KL’ and RC’, the training and test dataset sizes are 2,589 and 213 observations, respectively. The history length t of ME2* is fixed at 15 for red players and 13 for black players, which are optimal values determined using the existing data. To compare each estimated model’s KL divergence with that of the LSTM model on the new data, we performed paired t-tests. The null hypothesis states that there is no difference in KL divergence between the given model and the LSTM model, i.e., the mean difference is zero. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

distribution of cards is uniquely determined by $P_i^a/P_i^K = \exp(W_i^a - W_i^K)$ for $a = 1, 2, 3$, any model is characterized by its *normalized attractors* $\tilde{W}_i^a := W_i^a - W_i^K, a = 1, 2, 3$.

To formally assess the extent to which ME2* captures the regularities of data identified by LSTM, we conducted the following statistical test. We first regress a normalized LSTM attractor of number card j on a normalized ME2* attractor of card a ($a = 1, 2, 3$),

$$\tilde{W}^a_{i,\mathrm{LSTM}}(t) = \beta^a_{i,0} + \beta^a_{i,1}\tilde{W}^a_{i,\mathrm{ME2}^*}(t). \tag{9}$$

Using the OLS estimator of $(\hat{\beta}^a_{i,0}, \hat{\beta}^a_{i,1})$, we compute the residuals

$$R^a_i(t) := \tilde{W}^a_{i,\mathrm{LSTM}}(t) - \hat{\beta}^a_{i,0} - \hat{\beta}^a_{i,1}\tilde{W}^a_{i,\mathrm{ME2}^*}(t).$$

This represents the factors that are captured by LSTM but not captured by ME2*. To see how much predictive power this residual has, we estimate our **residual model**, whose attractor of card a is given by

$$W^a_i(t) := \begin{cases} \alpha^a_{i,0} + \alpha^a_{i,1}R^a_i(t) & \text{for } a = 1, 2, 3 \\ 0 & \text{for } a = K. \end{cases} \tag{10}$$

We compare the residual model and the constant (baseline) model in two ways.

First, we test if the cross validation outcomes of the residual model and the naive model are significantly different. For each round of CV, using the trained LSTM and ME2* models, we conduct OLS (9) and comupte the residual. Then, by minimizing the KL divergence of the residual model in the training data, we compute the value of parameters of the residual model (10). In the test data of the same round of CV, the residual is computed from the estimated parameters of LSTM and ME2* models and the coefficients of the OLS model (9), both computed in the training data of the same around. The residual thus obtained is plugged in to the residual model (10) whose parameters were estimated in the training data in the same round of CV. In this way, we compare the naive baseline model and the residual model in terms of the average KL divergence in the five rounds of CV.

Table 13 shows the results. The KL divergence of the residual models is 15% for the red player model and 9% for the black player model, and the difference in the prediction errors (KL) of the baseline and residual models is statistically significant according to the overall test.

Next, we conducted the model specification test using the trained LSTM and ME2* models with the *entire* dataset. We first estimate the residual model in Equation (10) and then conduct a likelihood ratio test with the null hypothesis $H_0\colon \alpha^1_{i,1} = \alpha^2_{i,1} = \alpha^3_{i,1} = 0$. The estimated parameters and the test results are shown in Table 14. The table shows that the predictive power of the residual model is significant.

These two residual tests show that, although ME2* achieves nearly the same predictive power as our best machine learning model (LSTM), it is not the case that they capture exactly the same mechanism. There are still certain regularities of human behavior captured by LSTM that remain to be decoded.

Table 13—Prediction Performance (Constant vs. Residual Model)

	Red Players			Black Players		
	#Params	KL	RC	#Params	KL	RC
Constant (Baseline)	3	1.360	0.000	3	1.354	0.000
Residual Model	6	1.356***	0.147	6	1.352***	0.091

Note: Average prediction performance of the selected models measured in the test data. Each performance score is the average of five train-test CV splits.

VII. Conclusion: Detect, Capture and Decode (DCD)

In summary, our paper suggests a possible way to utilize AI in economic research. The conventional approach has been to use our creative minds to detect the limitations of existing models and to think about how to improve them. This

Table 14—Estimated Parameters (Constant vs. Residual Model)

	Red Players		Black Players	
	Constant	Residual	Constant	Residual
α_{i0}^1	-0.392	-0.382	-0.425	-0.434
α_{i0}^2	-0.491	-0.491	-0.579	-0.597
α_{i0}^3	-0.568	-0.572	-0.588	-0.616
α_{i1}^1		0.924		1.046
α_{i1}^2		0.983		1.050
α_{i1}^3		0.977		1.079
Log Likelihood	-91525.8	-89865.2	-91153.1	-90372.0
#Parameters	3	6	3	6
LR statistic		3321.2***		1562.2***

Note: The LR Statistic is the likelihood ratio test statistic comparing the constant baseline model and the residual model. *p<0.1; **p<0.05; ***p<0.01.

paper suggests a procedure for semi-automatically doing some of those tasks, which we call Detect, Capture, and Decode (DCD). It proceeds in the following steps.

1) **Collect big data** that is an order of magnitude larger than the typical lab dataset. Our analysis suggests that this is necessary to capture hidden regularities, if any, using machine learning models, which have a large number of free parameters.

2) Compare the *external validities* of conventional models and machine learning models, by means of cross-validation, to detect if the existing models can be improved and to what extent. The notion of **relative completeness** offers a handy way to see how much predictive powers the conventional models have relative to the best machine learning model. Use **the overall test and the roundwise test** to see if the differences in the performances of the models are statistically significant.

3) If the existing models can be improved, certain regularities of human behavior that were not captured by the traditional models are captured by and encoded in the machine learning model(s). Try to **open the black box** of the better-performing machine learning model(s) and improve the existing models until the out-of-sample predictive error of the improved model is close to that of the best machine learning model.

4) **Decoding verification**: Examine if the improved model successfully decodes what is encoded in the best machine learning model, by means of the **residual tests**.

5) **Double-check the external validity** of the improved model using a separate dataset.

One of the challenges we face lies in the third step of this procedure, opening the black box of the machine learning models. This paper suggests the following method. Utilize interpretable machine learning models, such as the decision tree or LASSO model, along with complete black box models, such as deep learning models, and try to gain insights from the interpretable models. This method happened to worke well for the problem we considered in this paper, but it would be nice to come up with a more systematic way. Opening the black box of machine learning models is a hot research topic in AI, under the rubric of XAI (explainable AI). How Step 3 (opening the black box) might be automated is an important and promising future research topic.

References

Athey, Susan, and Guido W Imbens. 2019. "Machine Learning Methods that Economists Should Know About." *Annual Review of Economics*, 11: 685–725.

Augenblick, Ned, and Matthew Rabin. 2021. "Belief movement, uncertainty reduction, and rational updating." *The Quarterly Journal of Economics*, 136(2): 933–985.

Austern, Morgane, and Wenda Zhou. 2024. "Asymptotics of cross-validation." *Forthcoming in Annales de l'IHP Probabilités et statistiques.*

Bar-Hillel, Maya, and Willem A Wagenaar. 1991. "The perception of randomness." *Advances in Applied Mathematics*, 12(4): 428–454.

Brown, George W. 1951. "Iterative solution of games by fictitious play." *Act. Anal. Prod Allocation*, 13(1): 374.

Brown, James N, and Robert W Rosenthal. 1990. "Testing the minimax hypothesis: A re-examination of O'Neill's game experiment." *Econometrica*, 58(5): 1065–1081.

Camerer, Colin, and Teck Hua Ho. 1999. "Experience-weighted attraction learning in normal form games." *Econometrica*, 67(4): 827–874.

Camerer, Colin F. 2011. *Behavioral game theory: Experiments in strategic interaction.* Princeton University Press.

Camerer, Colin F. 2019. "Artificial intelligence and behavioral economics." *The Economics of Artificial Intelligence: An Agenda*, 587–608.

Camerer, Colin F, George Loewenstein, and Matthew Rabin. 2004. *Advances in behavioral economics.* Princeton University Press.

Chiappori, P-A, Steven Levitt, and Timothy Groseclose. 2002. "Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer." *American Economic Review*, 92(4): 1138–1151.

Erev, Ido, and Alvin E. Roth. 1998. "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria." *The American Economic Review*, 88(4): 848–881.

Fudenberg, Drew, and Annie Liang. 2019. "Predicting and understanding initial play." *American Economic Review*, 109(12): 4112–41.

Fudenberg, Drew, and Annie Liang. 2020. "Machine learning for evaluating and improving theories." *ACM SIGecom Exchanges*, 18(1): 4–11.

Fudenberg, Drew, and David K Levine. 1998. *The theory of learning in games.* Vol. 2, MIT press.

Fudenberg, Drew, Jon Kleinberg, Annie Liang, and Sendhil Mullainathan. 2021. "Measuring the Completeness of Economic Models." *Journal of Political Economy*, 130(4): 956–990.

Fudenberg, Drew, Wayne Gao, and Annie Liang. 2020. "How flexible is that functional form? quantifying the restrictiveness of theories." *arXiv preprint arXiv:2007.09213.*

Fudenberg, Drew, Wayne Gao, and Annie Liang. 2023. "How flexible is that functional form? Quantifying the restrictiveness of theories." *Forthcoming in Review of Economics and Statistics.*

Gauriot, Romain, Lionel Page, and John Wooders. 2023. "Expertise, gender, and equilibrium play." *Quantitative Economics*, 14(3): 981–1020.

Gers, Felix A, Jürgen Schmidhuber, and Fred Cummins. 2000. "Learning to forget: Continual prediction with LSTM." *Neural computation*, 12(10): 2451–2471.

Harrison, Glenn W, and John A List. 2004. "Field experiments." *Journal of Economic literature*, 42(4): 1009–1055.

Ke, Guolin, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. 2017. "Lightgbm: A highly efficient gradient boosting decision tree." *Advances in Neural Information Processing Systems*, 30.

Mullainathan, Sendhil, and Jann Spiess. 2017. "Machine learning: an applied econometric approach." *Journal of Economic Perspectives*, 31(2): 87–106.

Nunnari, Salvatore, Luca Congiu, and Sandrim Emiliano. 2022. "Laboratory Experiments in Top Economics Journals." *https://docs.google.com/spreadsheets/d/1434ApdJsvdtRNIpDR_03h-kmUffcL24_cJj7BXyCezs/edit#gid=491277435.*

Olah, Christopher. 2015. "Understanding LSTM Networks." https://colah.github.io/posts/2015-08-Understanding-LSTMs/.

O'Neill, Barry. 1987. "Nonmetric test of the minimax theory of two-person zerosum games." *Proceedings of the National Academy of Sciences*, 84(7): 2106–2109.

Palacios-Huerta, Ignacio. 2003. "Professionals play minimax." *The Review of Economic Studies*, 70(2): 395–415.

Peterson, Joshua C, David D Bourgin, Mayank Agrawal, Daniel Reichman, and Thomas L Griffiths. 2021. "Using large-scale experiments and machine learning to discover theories of human decision-making." *Science*, 372(6547): 1209–1214.

Peysakhovich, Alexander, and Jeffrey Naecker. 2017. "Using methods from machine learning to evaluate behavioral models of choice under risk and ambiguity." *Journal of Economic Behavior & Organization*, 133: 373–384.

Roth, Alvin E, and Ido Erev. 1995. "Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term." *Games and economic behavior*, 8(1): 164–212.

Tambe, Milind. 2011. *Security and game theory: algorithms, deployed systems, lessons learned.* Cambridge University Press.

Walker, Mark, and John Wooders. 2001. "Minimax play at Wimbledon." *American Economic Review*, 91(5): 1521–1538.

Appendix

A. Notes on the Overall Test

This section summarizes the details of the overall test introduced by Austern and Zhou (2024) and explained by Fudenberg, Gao and Liang (2020) discussed in Section VI.C. Although Fudenberg, Gao and Liang (2020) provides detailed instructions on how to use the overall test for economics practitioners, they cite an earlier version of Austern and Zhou (2024), where slightly different assumptions are used. Thus, following the notations in Fudenberg, Gao and Liang (2020), we provide instructions on how to use the overall test using the new assumptions.

Let $\mathcal{X}$ be a finite set of features, and let P_X be the marginal distribution of X. Let $\mathcal{Y} \subseteq \mathbb{R}^k$ be a compact set of outcomes. Let $\overline{\mathcal{F}} := \mathcal{Y}^{\mathcal{X}}$ be the set of all mappings from $\mathcal{X}$ to $\mathcal{Y}$, endowed with the usual topology. Each $f : \mathcal{X} \to \mathcal{Y}$ belonging to $\overline{\mathcal{F}}$ is called a *prediction rule*. For a compact set Θ of parameters, we define a parametric model, denoted by $\mathcal{F}_\Theta = \{f_\theta\}_{\theta\in\Theta} \subseteq \overline{\mathcal{F}}$, where each f_θ is continuous in θ.[25] A loss function is given by $l : \overline{\mathcal{F}} \times \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$, where $l\left(f, (x, y)\right)$ is the loss between the prediction $f\left(x\right)$ and the true outcome y. Let $Z_i = (X_i, Y_i)$ be the observation i.[26] The researchers observe data denoted by $\{Z_i\}_{i=1}^N$, where each observation is independently and identically distributed according to the distribution P.

Given the data set $\{Z_i\}_{i=1}^N$, we can compute the out-of-sample prediction error of a parametric model $\mathcal{F}_\Theta$ as follows:

1) Divide the data $\{Z_i\}_{i=1}^N$ into K groups, denoted by $\{G_k\}_{k=1}^K$, where each G_k has approximately the same size. For simplicity, assume that N/K is an integer so that each G_k has N/K observations.
2) In each k-th process of the cross-validation, $\{Z_i\}_{i=1}^N \setminus G_k$ is used as the training data and G_k is used as the test data.
3) For each $k \in \{1, \dots, K\}$, compute the best prediction rule that minimizes the average sample error for prediction of the k-th training data $\{Z_i\}_{i=1}^N \setminus G_k$:

$$\hat{f}_{\mathcal{F}_\Theta,CV}^{-k} := \arg\min_{f\in\mathcal{F}_\Theta} \frac{1}{N - N/K} \sum_{Z_j \in \{Z_i\}_{i=1}^N \setminus G_k} l\left(f, Z_j\right).$$

   This best prediction rule gives the average out-of-sample error for the prediction of the k-th test data G_k:

$$\hat{e}_{k,\mathcal{F}_\Theta,CV} := \frac{1}{N/K} \sum_{Z_j \in G_k} l\left(\hat{f}_{\mathcal{F}_\Theta,CV}^{-k}, Z_j\right).$$

[25] Since $\mathcal{X}$ is finite, we can regard $\Theta \subseteq \mathbb{R}^p$, where $p < \infty$.
[26] In our setting, each observation corresponds to the choices of a player for 30 periods.

4) Finally, the average out-of-sample error across cross-validation is as follows:

$$\hat{e}_{\mathcal{F}_\Theta,CV} \coloneqq \frac{1}{K}\sum_{k=1}^{k} \hat{e}_{k,\mathcal{F}_\Theta,CV}.$$

For given data $\{Z_i\}_{i=1}^{N}$ and a given parametric model $\mathcal{F}_\Theta$, we define the conditional out-of-sample error by

$$e_{\mathcal{F}_\Theta} \coloneqq \mathbb{E}_{Z_{N+1}\sim P}\left[l\left(\hat{f}_{\mathcal{F}_\Theta}, Z_{N+1}\right)\right],$$

where $\hat{f}_{\mathcal{F}_\Theta}$ is the best prediction rule that minimizes the average sample error for prediction of the entire training data $\{Z_i\}_{i=1}^{N}$. That is,

$$\hat{f}_{\mathcal{F}_\Theta} \coloneqq \arg\min_{f\in\mathcal{F}_\Theta} \frac{1}{N}\sum_{Z_j\in\{Z_i\}_{i=1}^{N}} l\left(f, Z_j\right).$$

As data $\{Z_i\}_{i=1}^{N}$ is realized according to the distribution P^N, we can define the unconditional out-of-sample error by

$$e_{\mathcal{F}_\Theta,N} \coloneqq \mathbb{E}_{\{Z_i\}_{i=1}^{N}\sim P^N}\left[e_{\mathcal{F}_\Theta}\right].$$

By construction of $\hat{e}_{\mathcal{F}_\Theta,CV}$, $\mathbb{E}_{\{Z_i\}_{i=1}^{N}\sim P^N}\left[\hat{e}_{\mathcal{F}_\Theta,CV}\right] = e_{\mathcal{F}_\Theta,\frac{K-1}{K}N}$ holds. That is, the expected average out-of-sample errors across cross-validation given N observations is equal to the unconditional out-of-sample error given $\frac{K-1}{K}N$ observations.

Using this property, we consider the asymptotics of $\hat{e}_{\mathcal{F}_\Theta,CV} - e_{\mathcal{F}_\Theta,\frac{K-1}{K}N}$. To this end, we introduce additional assumptions. For a generic prediction rule $f_\theta \in \mathcal{F}_\Theta$, we define

$$l_\Theta\left(\theta, Z_i\right) \coloneqq l\left(f_\theta, Z_i\right).$$

Also, we define the best prediction rule f_{θ^*} by

$$\theta^* \in \arg\min_{\theta\in\Theta} \mathbb{E}_{Z\sim P}\left[l_\Theta\left(\theta, Z\right)\right].$$

ASSUMPTION 1 (Austern and Zhou (2024)): *A model $\mathcal{F}_\Theta$ and its best prediction rule f_{θ^*} satisfy the following:*

1) $l_\Theta\left(\theta, Z_i\right)$ is three times differentiable and strictly convex in θ.

2) $\max_{k\in\{1,\dots,p\}}\left\|\frac{\partial l_\Theta(\theta^*,Z_i)}{\partial\theta_k}\right\|_{L50} < \infty$, and $\sup_{N'\in\mathbb{N}}\mathbb{E}_{\left(Z,\{Z_i\}_{i=1}^{N'}\right)\sim P^{N'+1}}\left[l_\Theta^8\left(\hat{\theta}_{N'}, Z\right)\right] < \infty$, where $\hat{\theta}_{N'}$ satisfies $f_{\hat{\theta}_{N'}} = \hat{f}_{\mathcal{F}_\Theta}$, where $\hat{f}_{\mathcal{F}_\Theta}$ is derived using $\{Z_i\}_{i=1}^{N'}$.

3) *There is a convex open neighborhood $\mathcal{O}_{\theta^*}$ of θ^* such that*
   a) $\| \sup_{\theta \in \mathcal{O}_{\theta^*}} \|\nabla_\theta l_\Theta (\theta, Z_i)\|_2 \|_{L32} < \infty$,
   b) $\| \sup_{\theta \in \mathcal{O}_{\theta^*}} \lambda_{max} \left(\nabla^2_\theta l_\Theta (\theta, Z_i) \right) \|_{L16} < \infty$, *where λ_{max} is a mapping from a matrix to its highest eigenvalue,*
   c) $\max_{k \in \{1,\dots,p\}} \left\| \sup_{\theta \in \mathcal{O}_{\theta^*}} \|\nabla^3_\theta l_\Theta (\theta, Z_i)_{k,\cdot,\cdot} \|_{op} \right\|_{L32} < \infty$, *where $\|\cdot\|_{op}$ is the operator 2-norm, and*
   d) *there exists $c > 0$ such that $\inf_{\theta \in \mathcal{O}_{\theta^*}} \lambda_{min} \left(\nabla^2_\theta l_\Theta (\theta, Z_i) \right) \geq c$ holds almost surely, where λ_{min} is a mapping from a matrix to its lowest eigenvalue.*

THEOREM 1 (Proposition 5 in Austern and Zhou (2024)): *Under Assumption 1, we have*

$$\sqrt{N} \left(\hat{e}_{\mathcal{F}_\Theta, CV} - e_{\mathcal{F}_\Theta, \frac{K-1}{K} N} \right) \xrightarrow{d} \mathcal{N} \left(0, \, Var(l(f_{\theta^*}, Z_i))\right).$$

This theorem provides the foundation for the overall test. Note that the following estimator consistently estimates the asymptotic variance of $\hat{e}_{\mathcal{F}_\Theta, CV}$:

$$\hat{\sigma}^2_{\mathcal{F}_\Theta} := \frac{1}{K} \sum_{k=1}^{K} \frac{1}{\frac{N}{K} - 1} \sum_{Z_j \in G_k} \left[l\left(\hat{f}^{-k}_{\mathcal{F}_\Theta, CV}, Z_j \right) - \frac{1}{\frac{N}{K}} \sum_{Z_{j'} \in G_k} l\left(\hat{f}^{-k}_{\mathcal{F}_\Theta, CV}, Z_{j'} \right) \right]^2.$$

PROPOSITION 1 (Proposition 1 in Austern and Zhou (2024)): *Under Assumption 1, we have*

$$\hat{\sigma}^2_{\mathcal{F}_\Theta} \xrightarrow{p} Var(l(f_{\theta^*}, Z_i)).$$

A.1. Hypothesis Testing

We are interested in testing whether the unconditional out-of-sample error for a model $\mathcal{F}_\Theta$ is significantly smaller than that for another model $\mathcal{F}_{\Theta'}$. We define

$$\Delta l \left(f_{\theta^*}, f_{\theta'^*}, Z_i \right) := l_\Theta (\theta^*, Z_i) - l_{\Theta'} \left({\theta'}^*, Z_i \right).$$

Theorem 1 implies the following result:

PROPOSITION 2: *If Assumption 1 holds for both $\mathcal{F}_\Theta$ and $\mathcal{F}_{\Theta'}$, then we have*

$$\sqrt{N} \left[\left(\hat{e}_{\mathcal{F}_\Theta, CV} - \hat{e}_{\mathcal{F}_{\Theta'}, CV} \right) - \left(e_{\mathcal{F}_\Theta, \frac{K-1}{K} N} - e_{\mathcal{F}_{\Theta'}, \frac{K-1}{K} N} \right) \right] \xrightarrow{d} \mathcal{N} \left(0, \, Var\left(\Delta l \left(f_{\theta^*}, f_{\theta'^*}, Z_i \right) \right) \right).$$

PROOF:
The proof of Lemma C.1. in Fudenberg, Gao and Liang (2023), where the result of Theorem 1 is used, applies when we replace their $\overline{\mathcal{F}}$ with our $\mathcal{F}_{\Theta'}$.

Note that, as in Proposition 1, we can construct a consistent estimator of the asymptotic variance of $\hat{e}_{\mathcal{F}_\Theta,CV} - \hat{e}_{\mathcal{F}_{\Theta'},CV}$:

$$
\begin{aligned}
\hat{\sigma}^2_{\Theta+\Theta'} :=& 2\hat{\sigma}^2_{\mathcal{F}_\Theta} + 2\hat{\sigma}^2_{\mathcal{F}_{\Theta'}} - \hat{\sigma}^2_{\mathcal{F}_\Theta+\mathcal{F}_{\Theta'}} \\
=& \frac{1}{K}\sum_{k=1}^{K} \frac{1}{\frac{N}{K}-1} \sum_{Z_j \in G_k} \Bigg[l\left(\hat{f}^{-k}_{\mathcal{F}_\Theta,CV}, Z_j\right) - \frac{1}{\frac{N}{K}} \sum_{Z_{j'} \in G_k} l\left(\hat{f}^{-k}_{\mathcal{F}_\Theta,CV}, Z_{j'}\right) \\
& + l\left(\hat{f}^{-k}_{\mathcal{F}_{\Theta'},CV}, Z_j\right) - \frac{1}{\frac{N}{K}} \sum_{Z_{j'} \in G_k} l\left(\hat{f}^{-k}_{\mathcal{F}_{\Theta'},CV}, Z_{j'}\right) \Bigg]^2 .
\end{aligned}
$$

PROPOSITION 3: *Under Assumption 1, we have*

$$
\hat{\sigma}^2_{\Theta+\Theta'} \xrightarrow{p} Var\left(\Delta l\left(f_{\theta^*}, f_{\theta'^*}, Z_i\right)\right).
$$

PROOF:

The proof of Lemma C.2. in Fudenberg, Gao and Liang (2023), where the result of Proposition 1 is used, applies when we replace their $\overline{\mathcal{F}}$ with our $\mathcal{F}_{\Theta'}$.

Therefore, the standard error of $\hat{e}_{\mathcal{F}_\Theta,CV} - \hat{e}_{\mathcal{F}_{\Theta'},CV}$ can be derived by

$$
SE\left(\hat{e}_{\mathcal{F}_\Theta,CV} - \hat{e}_{\mathcal{F}_{\Theta'},CV}\right) := \sqrt{\hat{\sigma}^2_{\Theta+\Theta'}/N}.
$$

B. Details of Machine Learning Models

B.1. List of Variables for LASSO, Decision Tree, and LightGBM

We employed a total of 2,448 candidate variables for LASSO, decision tree, and LightGBM listed in Table B1.[27] These variables were selected according to the following criteria. First, we include all dummies of actions and action profiles taken in $t-1$, $t-2$, $t-3$, and $t-4$. Second, we include one- and two-period history dummies. In addition, since we do not have enough observations of longer histories, we focus on the reduced histories, such as histories of K-number profiles and histories of win-lose patterns, to incorporate the associated dummies in the last four periods.

Table B1—The list of variables included in ML models

List of (Dummy) Variables
Constant
Period dummy (t=5, 6, . . . , 30)
R's action in t-1, t-2, t-3, t-4
B's action in t-1, t-2, t-3, t-4
Action profile in t-1, t-2, t-3, t-4
R's action was a number or K in t-1, t-2, t-3, t-4
B's action was a number or K in t-1, t-2, t-3, t-4
Action profile (number or K) in t-1, t-2, t-3, t-4
History of R's actions in the last 2, 3, 4 periods
History of B's actions in the last 2, 3, 4 periods
History of action profiles in the last 2 periods
History of action profiles (number or K) in the last 2, 3, 4 periods
R played 1 (or 2, 3, K) consecutively in the last three periods
B played 1 (or 2, 3, K) consecutively in the last three periods
R did not play 1 (or 2, 3, K) in the last three periods
B did not play 1 (or 2, 3, K) in the last three periods
R won or lost in periods t-1, t-2, t-3, t-4
History of winners in the last 2, 3, 4 periods
History of R's actions (number or K) and winners in t-1, t-2, t-3, t-4
History of B's actions (number or K) and winners in t-1, t-2, t-3, t-4
History of action profiles (number or K) and winners in t-1, t-2, t-3, t-4
R played K (or numbers) in t-1, ..., t-n consecutively (n=4, 5, ..., 29)
B played K (or numbers) in t-1, ..., t-n consecutively (n=4, 5, ..., 29)

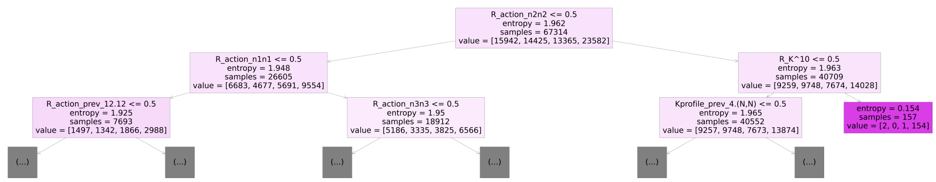
Notes: The list of variables included in the LASSO, decision tree, and LightGBM models. Here, t is the current period; a model predicts the action in period t using the variables listed above.

[27]The number of parameters we initially created was 2,702. Dropping variables whose values were exactly the same as another variable then left us with 2,448 variables.

B.2. Illustration of Decision Tree

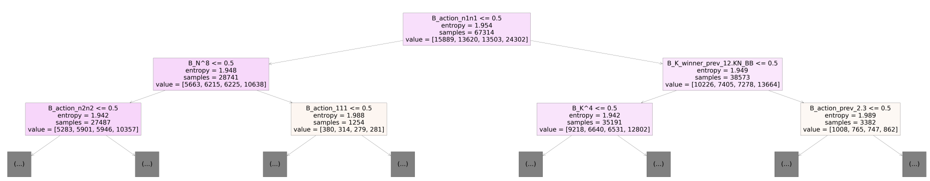
In this section, we show the sample of decision trees for the red player (CV-1). Each node, except for the terminal nodes, corresponds to a feature, a property of three-period history whose occurrences can be answered by Yes or No. If the answer to the question is Yes, we go to the bottom right, and if it is No, we go to the bottom left.

FIGURE B1. PART OF THE DECISION TREE FOR THE RED PLAYER TRAINED BY ALL DATA.



Notes: In each node, the first line represents the feature that partitions data, the second line represents the entropy in the node, the third line represents the number of observations in the node, and the last line represents how many times each action is chosen (from left to right: 1, 2, 3, K). Note that choosing a feature for splitting based on the entropy is equivalent to choosing a feature for splitting based on the KL divergence. R_action_nana asks if the red player did not play $a \in \{1, 2, 3, K\}$ in the last two periods, $R\_action\_prev\_12.12$ asks if the red player played 1 at t-1 and 2 at t-2, R_K^{10} asks if the red player played K in the last 10 periods consecutively, and $Kprofile\_prec\_4.(N, N)$ asks if a pair of numbers were player at t-4.

FIGURE B2. PART OF THE DECISION TREE FOR THE BLACK PLAYER TRAINED BY ALL DATA.



Notes: In each node (except for the node on the far right), the first line represents the feature that partitions data, the second line represents the entropy in the node, the third line represents the number of observations in the node, and the last line represents how many times each action is chosen (from left to right: 1, 2, 3, K). Note that choosing a feature for splitting based on the entropy is equivalent to choosing a feature for splitting based on the KL divergence. B_action_nana asks if the black player did not play $a \in \{1, 2, 3, K\}$ in the last two periods, $B\_N\hat{}n$ asks if the black player played numbers in the last n periods consecutively, B_action_111 asks if the black player played 1 in the last periods consecutively, $B\_K\_winner\_prev\_12.KN\_BB$ asks if the black player won by K at t-1 and won by numbers at t-2, and $B\_action\_prev\_2.3$ asks if the black player played 3 at t-2.

B.3. Parameters that LASSO Commonly Selected across All Five Cross-validation Splits

In this section, we summarize in Table B2 to B7 the list of coefficients that appear in all five CVs. Since we already showed the table for the red player, card

1 (Table 5) and for the red player, card K (Table 6) in Section IV.B, we do not repeat them here.

Table B2—Commonly Selected Variables in LASSO (Red player, Card 2)

	$\beta_{\mathrm{R},2}$	
	Value	Count
Period dummy (t=8)	−0.049	2589
Period dummy (t=11)	0.045	2589
R played 2 at t-1	−0.411	14494
B played 2 at t-1	−0.034	13675
Action profile at t-1 was (K,1)	0.080	5386
Action profile at t-1 was (K,2)	−0.053	4284
R played 2 at t-2	−0.192	14470
Action profile at t-2 was (2,2)	−0.026	3157
Action profile at t-3 was (2,2)	−0.034	3140
Action profile at t-4 was (2,1)	0.096	3415
R played 2 at t-1, 2 at t-2	0.393	2359
R played 1 at t-1, 3 at t-2, K at t-3	0.062	1353
R played 2 at t-1, 2 at t-2, 2 at t-3	0.510	498
R played K at t-1, a number at t-2, a number at t-3, K at t-4	0.058	3982
B played K at t-1, 1 at t-2, 2 at t-3	−0.108	1355
R won by playing 2 at t-1	−0.055	6411
R lost by playing 2 at t-3	−0.021	8039
R lost by playing 1 at t-1, lost by playing 2 at t-2	−0.142	1254
R lost by playing K at t-1, lost by playing 3 at t-2	0.071	1659
B won by playing 2 at t-4	−0.073	7430
B won by playing a number at t-1, lost by playing a number at t-2	−0.034	6817

Notes: The Value column presents the estimated coefficients based on the entire dataset. If LASSO with the entire dataset excludes any variable above during the estimation, the corresponding Value entry is left blank. The Count column specifies the number of histories in which each dummy variable is assigned a value of 1.

Table B3—Commonly Selected Variables in LASSO (Red player, Card 3)

	$\beta_{\mathrm{R},3}$	
	Value	Count
Constant	−0.053	67314
Period dummy (t=11)	−0.062	2589
Period dummy (t=18)	0.077	2589
R played 3 at t-1	−0.394	13449
Action profile at t-1 was (2,1)	0.049	3464
R played 3 at t-2	−0.173	13488
B played 3 at t-2	−0.066	13521
2-period K-history is ((N,N),(N,N))	−0.080	12206
3-period K-history is ((N,K),(N,N),(N,N))	−0.072	2566
R played 2 at t-1, 1 at t-2	0.043	3883
R played 3 at t-1, 3 at t-2	0.331	1976
R played 1 at t-1, 2 at t-2, K at t-3	0.196	1464
R played 2 at t-1, 1 at t-2, K at t-3	0.149	1506
R played 3 at t-1, 3 at t-2, 3 at t-3	0.283	364
R played 3 at t-1, K at t-2, 1 at t-3	−0.199	1325
R played K at t-1, K at t-2, 3 at t-3	−0.120	1476
R played a number at t-1, K at t-2, a number at t-3	0.010	10752
B played K at t-1, K at t-2, 3 at t-3	−0.077	1561
Winners at t-2 was R	−0.020	28161
Winners at t-1, t-2, t-3 are R, B, R	−0.051	6858
R won by playing 1 at t-4	−0.067	6656
R lost by playing 3 at t-1, lost by playing 3 at t-2	0.144	591
B won by playing K at t-1, lost by playing 1 at t-2	−0.066	1436
R consecutively played a number in the last 8 periods	0.110	1967

Notes: The Value column presents the estimated coefficients based on the entire dataset. If LASSO with the entire dataset excludes any variable above during the estimation, the corresponding Value entry is left blank. The Count column specifies the number of histories in which each dummy variable is assigned a value of 1.

Table B4—Commonly Selected Variables in LASSO (Black player, Card 1)

	$\beta_{\mathrm{B},1}$	
	Value	Count
Period dummy (t=10)	0.107	2589
Period dummy (t=16)	−0.071	2589
Period dummy (t=19)	−0.034	2589
R played 1 at t-1	−0.139	15988
B played 1 at t-1	−0.442	15832
B played 3 at t-1	0.034	13517
Action profile at t-1 was (1,1)	−0.070	3978
Action profile at t-1 was (3,2)	0.080	2778
B played 1 at t-2	−0.170	15792
Action profile at t-2 was (3,1)	−0.118	3007
Action profile at t-2 was (K,3)	−0.056	4380
R played K at t-3	0.037	23380
Action profile at t-3 was (K,1)	0.033	5366
B played K at t-4	0.016	24480
Action profile at t-4 was (K,1)	0.052	5450
Action profile at t-4 was (K,2)	−0.030	4315
2-period K-history is ((N,N),(N,N))	0.021	12206
4-period K-history is ((N,N),(N,N),(N,K),(N,N))	−0.109	1227
R played 2 at t-1, 3 at t-2	0.116	3116
R played 2 at t-1, K at t-2	0.036	5136
R played 3 at t-1, 2 at t-2	0.068	3215
R played 3 at t-1, 1 at t-2, 2 at t-3	0.122	920
R played 3 at t-1, 2 at t-2, 1 at t-3	0.068	956
R played K at t-1, a number at t-2, a number at t-3, a number at t-4	0.039	6600
B played 1 at t-1, 1 at t-2	0.410	2883
B played 1 at t-1, K at t-2	−0.072	6011
B played K at t-1, 1 at t-2	−0.110	5985
B played 1 at t-1, 1 at t-2, 1 at t-3	0.141	718
B played 1 at t-1, K at t-2, 3 at t-3	−0.114	1363
B played K at t-1, 2 at t-2, 3 at t-3	0.068	1181
B played 1 at t-1, 1 at t-2, 1 at t-3, 1 at t-4	0.945	254
B played 1 at t-1, 1 at t-2, K at t-3, K at t-4	0.147	402
B played K at t-1, K at t-2, a number at t-3, K at t-4	−0.042	1937
Winners at t-1, t-2, t-3 are B, R, R	0.077	6808
R lost by playing 1 at t-2	−0.046	9356
R lost by playing 2 at t-4	0.026	8012
B won by playing 1 at t-1, won by playing 1 at t-2	0.193	989
K-profile and winner were (N,N) and B at t-1, (N,N) and B won at t-2, (N,N) and B won at t-3	0.407	300

Notes: The Value column presents the estimated coefficients based on the entire dataset. If LASSO with the entire dataset excludes any variable above during the estimation, the corresponding Value entry is left blank. The Count column specifies the number of histories in which each dummy variable is assigned a value of 1.

Table B5—Commonly Selected Variables in LASSO (Black player, Card 2)

	$\beta_{\mathrm{B},2}$	
	Value	Count
Period dummy (t=8)	−0.109	2589
Period dummy (t=12)	−0.094	2589
Period dummy (t=17)	−0.051	2589
R played 2 at t-1	−0.149	14494
B played 2 at t-1	−0.328	13675
R played 2 at t-2	−0.041	14470
B played 2 at t-2	−0.068	13645
B played K at t-3	−0.015	24427
R played 3 at t-4	−0.046	13425
R played K at t-4	−0.055	23435
B played K at t-4	−0.048	24480
Action profile at t-4 was (1,2)	0.037	3467
Action profile at t-4 was (1,3)	−0.063	3189
Action profile at t-4 was (K,2)	0.037	4315
R played K at t-1, a number at t-2, K at t-3, K at t-4	−0.040	1644
B played 2 at t-1, 2 at t-2	0.260	2140
B played 1 at t-1, K at t-2, 3 at t-3	0.049	1363
B played 1 at t-1, K at t-2, K at t-3	−0.026	2027
B played 2 at t-1, 2 at t-2, 2 at t-3	0.210	429
B played 3 at t-1, 1 at t-2, K at t-3	0.099	1376
B played K at t-1, 3 at t-2, 1 at t-3	0.144	1314
B played 2 at t-1, 2 at t-2, 2 at t-3, 2 at t-4	0.694	130
B played K at t-1, K at t-2, K at t-3, K at t-4	−0.135	1497
B played K at t-1, K at t-2, a number at t-3, a number at t-4	−0.056	3170
Winners at t-1, t-2, t-3 are R, R, B	−0.033	6852
Winners at t-1, t-2, t-3, t-4 are R, B, B, R	−0.035	3972
R lost by playing 1 at t-3	0.047	9411
R lost by playing 1 at t-4	0.001	9450
R won by playing 2 at t-4	0.014	6336
R won by playing 3 at t-4	−0.061	5731
R won by playing 2 at t-1, won by playing K at t-2	−0.190	910
B won by playing 3 at t-1	0.073	7364
B won by playing 3 at t-1, won by playing 1 at t-2	0.106	1084
B consecutively played a number in the last 6 periods	0.077	4447

Notes: The Value column presents the estimated coefficients based on the entire dataset. If LASSO with the entire dataset excludes any variable above during the estimation, the corresponding Value entry is left blank. The Count column specifies the number of histories in which each dummy variable is assigned a value of 1.

Table B6—Commonly Selected Variables in LASSO (Black player, Card 3)

	$\beta_{B,3}$	
	Value	Count
Constant	−0.033	67314
Period dummy (t=6)	−0.101	2589
Period dummy (t=19)	0.027	2589
Period dummy (t=23)	−0.044	2589
Period dummy (t=27)	0.023	2589
R played 3 at t-1	−0.080	13449
B played 3 at t-1	−0.172	13517
Action profile at t-1 was (3,3)	−0.181	2971
B played 3 at t-2	−0.186	13521
Action profile at t-2 was (3,3)	−0.057	2989
Action profile at t-3 was (3,3)	−0.138	2970
B played 1 at t-4	−0.050	15877
B played K at t-4	−0.014	24480
Action profile at t-4 was (2,2)	0.049	3115
Action profile at t-4 was (2,K)	−0.020	4897
4-period K-history is ((K,N),(N,N),(N,N),(N,N))	−0.089	1018
R played a number at t-1, K at t-2, K at t-3, a number at t-4	−0.035	3239
B played 2 at t-1, 1 at t-2	0.121	3501
B played 3 at t-1, 1 at t-2	−0.163	3423
B played 3 at t-1, 3 at t-2	0.079	2055
B played 3 at t-1, K at t-2	−0.159	4966
B played 1 at t-1, 2 at t-2, K at t-3	0.102	1319
B played 2 at t-1, K at t-2, 1 at t-3	0.168	1360
B played 3 at t-1, 3 at t-2, 3 at t-3	0.550	382
B played 3 at t-1, K at t-2, 2 at t-3	−0.141	1145
B played K at t-1, K at t-2, 2 at t-3	0.081	1563
B played K at t-1, K at t-2, K at t-3	−0.056	3287
Winners at t-1 was R	−0.024	28165
Winners at t-2 was R	−0.013	28161
Winners at t-1, t-2, t-3 are R, B, R	−0.035	6858
R lost by playing 3 at t-2	−0.085	7717
R lost by playing 3 at t-1, lost by playing 2 at t-2	−0.130	993
R lost by playing K at t-1, lost by playing 3 at t-2	0.087	1659
B lost by playing 3 at t-1, won by playing 2 at t-2	−0.165	804
B won by playing 3 at t-1, won by playing 3 at t-2	0.227	547
R consecutively played a number in the last 7 periods	−0.060	2879

Notes: The Value column presents the estimated coefficients based on the entire dataset. If LASSO with the entire dataset excludes any variable above during the estimation, the corresponding Value entry is left blank. The Count column specifies the number of histories in which each dummy variable is assigned a value of 1.

Table B7—Commonly Selected Variables in LASSO (Black player, Card K)

	$\beta_{\mathrm{B},K}$	
	Value	Count
Constant	0.147	67314
Period dummy (t=5)	0.132	2589
Period dummy (t=20)	−0.057	2589
Period dummy (t=30)	0.092	2589
R played K at t-1	0.089	23383
Action profile at t-2 was (1,3)	0.062	3195
B played 2 at t-3	−0.013	13618
B played K at t-3	0.087	24427
K-profile at t-3 was (N,K)	0.054	15057
B played K at t-4	0.158	24480
2-period K-history is ((N,K),(K,K))	0.069	2295
2-period K-history is ((N,N),(K,N))	0.059	5873
3-period K-history is ((K,K),(N,K),(N,N))	−0.101	772
3-period K-history is ((K,K),(N,N),(N,N))	−0.099	1789
3-period K-history is ((K,N),(N,N),(N,N))	0.147	2639
3-period K-history is ((N,N),(K,N),(K,K))	0.139	769
4-period K-history is ((N,N),(N,K),(K,K),(N,N))	0.120	498
4-period K-history is ((N,N),(N,N),(N,N),(N,N))	−0.074	2352
R played 1 at t-1, K at t-2, 3 at t-3	−0.085	1305
R played K at t-1, a number at t-2, K at t-3	0.076	5009
B played K at t-1, K at t-2	0.136	8394
B played 1 at t-1, 2 at t-2, 3 at t-3	0.085	844
B played 3 at t-1, 2 at t-2, 1 at t-3	0.172	942
B played K at t-1, 1 at t-2, 1 at t-3	0.125	978
B played K at t-1, K at t-2, K at t-3	0.072	3287
B played a number at t-1, a number at t-2	0.058	27062
B played K at t-1, a number at t-2, K at t-3, K at t-4	0.124	1955
B played K at t-1, a number at t-2, K at t-3, a number at t-4	−0.126	3738
B played a number at t-1, K at t-2, a number at t-3, K at t-4	0.112	3744
B played a number at t-1, a number at t-2, a number at t-3, a number at t-4	0.130	10619
Winner at t-1 was R	0.038	28165
R lost by playing 2 at t-1, lost by playing 1 at t-2	0.046	1278
B won by playing K at t-1, won by playing a number at t-2	−0.202	5619
B won by playing a number at t-1, lost by playing K at t-2	−0.105	3092
B won by playing a number at t-1, won by playing a number at t-2	0.060	8485
K-profile and winner were (N,N) and B at t-1	−0.058	10109
K-profile and winner were (N,N) and B at t-1	−0.054	10092
K-profile and winner were (N,K) and B at t-1, (K,K) and R won at t-2, (N,N) and R won at t-3	0.102	780
K-profile and winner were (N,N) and B at t-1	0.000	10046
B consecutively played K in the last 5 periods	0.159	735
B consecutively played K in the last 7 periods	0.447	232
B consecutively played a number in the last 6 periods	−0.058	4447
B consecutively played a number in the last 8 periods	−0.172	2175
B consecutively played a number in the last 9 periods	−0.148	1612
B consecutively played a number in the last 11 periods	−0.281	954
B consecutively played a number in the last 13 periods	−0.199	616

Notes: The 'Value' column presents the estimated coefficients based on the entire dataset. If LASSO with the entire dataset excludes any variable above during the estimation, the corresponding 'Value' entry is left blank. The 'Count' column specifies the number of histories in which each dummy variable is assigned a value of 1.

C. Additional Performance Comparison

In Appendix C, we provide more detailed performance comparisons that were omitted in the main text. Appendix C.1 enumerates the performance metrics for all the models discussed in the paper. We also include additional results from new models to assess robustness. Appendix C.2 presents the performance scores of the selected models (as same as Table 8 and 9 in the main text) using alternative criteria. Here, we calculate (i) L1 and L2 losses using the test data and (ii) traditional econometric measures of model selection, AIC and BIC, based on the training data. We will compare those measures with KL divergence and strategic error rates.

C.1. Omitted Model Performances

Table C8 presents the performance metrics (KL divergence, strategic error rates, and relative completeness) for all the models we estimated. We computed the scores in exactly the same way as outlined in the main text — we partitioned the data into five disjoint groups, trained each model on four of these groups, evaluated errors using the remaining group, and then averaged the scores across the five cross-validation splits.

For the sake of robustness, we estimated two additional models: (i) Serial Correlation of Order 4 and (ii) ME2, both without the constraint of identical coefficients for the number cards. The respective lines for Serial Correlation (Order 4, no restriction) and ME2 (no restriction) correspond to those models. As noted earlier, these unrestricted versions exhibit marginally lower performance than their restricted counterparts.[28] For instance, the relative completeness (RC) of the unrestricted ME2 for the red player is 0.852, slightly lower than the 0.870 RC for the restricted version.

In addition, we include the models that mix EWA and Serial Correlation. The lines for EWA+Serial Correlation (Order $1, 4$) in Table C8 correspond to those models. We can see that the RC of EWA+Serial Correlation (Order 4) is 0.628 for the red player and 0.587 for the black player, which is better than the original EWA (0.276 and 0.195) and Serial Correlation or Order 4 (0.293 and 0.334) but worse than ME1 (0.825 and 0.753). This result indicates that our specific form of serial correlation—(1) whether the card was chosen consecutively in the last t periods and (2) whether the card was not chosen in the last t periods—explains a greater portion of real human behavior.

[28] Except for KL and RC of ME2 for the black player.

Table C8—Prediction Performance Comparison (All Models)

	Red Players				Black Players			
	#Params	KL	SER	RC	#Params	KL	SER	RC
Conventional Models								
Constant (Baseline)	3	1.360	0.350	0.000	3	1.354	0.562	0.000
1-period Nonparametric	48	1.354	0.351	0.202	48	1.347	0.540	0.281
2-period Nonparametric	768	1.354	0.341	0.215	768	1.347	0.539	0.297
1-period KN history	12	1.360	0.350	0.003	12	1.353	0.562	0.052
2-period KN history	48	1.357	0.347	0.082	48	1.352	0.559	0.099
3-period KN history	192	1.355	0.337	0.155	192	1.352	0.558	0.102
4-period KN history	768	1.357	0.336	0.083	768	1.355	0.560	−0.036
RL	4	1.359	0.350	0.036	4	1.352	0.559	0.088
BL	4	1.354	0.335	0.213	4	1.360	0.563	−0.244
EWA	8	1.352	0.332	0.276	8	1.349	0.559	0.195
Nested EWA	9	1.351	0.331	0.310	9	1.349	0.560	0.184
Serial Correlation								
Order 1	5	1.354	0.350	0.204	5	1.349	0.541	0.199
Order 2	7	1.353	0.348	0.235	7	1.348	0.537	0.235
Order 3	9	1.353	0.348	0.251	9	1.347	0.537	0.268
Order 4	11	1.351	0.343	0.293	11	1.346	0.536	0.334
Order 5	13	1.351	0.339	0.322	13	1.345	0.532	0.367
Order 6	15	1.349	0.336	0.359	15	1.344	0.527	0.411
Order 4, no restriction	19	1.351	0.344	0.291	19	1.346	0.536	0.330
Machine Learning Models								
LASSO	408.2	1.344	0.324	0.562	476.0	1.339	0.526	0.589
Decision Tree	40.4	1.352	0.343	0.273	41.6	1.348	0.536	0.248
LightGBM	30.0	1.344	0.325	0.547	30.0	1.340	0.527	0.556
DNN	6208.0	1.347	0.332	0.455	6428.0	1.343	0.534	0.432
LSTM	5524.0	1.331	0.308	1.000	11204.0	1.329	0.507	1.000
Improved Models								
ME1 ($t = 4$)	24	1.336	0.314	0.825	24	1.335	0.521	0.753
ME2 ($t = 4$)	38	1.335	0.311	0.870	38	1.333	0.517	0.848
ME2 ($t = 4$, no restriction)	68	1.335	0.312	0.852	68	1.332	0.518	0.852
ME2*	126	1.332	0.311	0.961	110	1.330	0.516	0.932
ME2* (no restriction)	244	1.337	0.312	0.792	212	1.332	0.517	0.854
EWA + Serial Correlation								
Order 1	11	1.344	0.327	0.536	11	1.342	0.528	0.486
Order 4	17	1.342	0.321	0.628	17	1.339	0.522	0.587
Human			0.419			0.581		

Notes: Average prediction performance measured in the test data. KL, SER, and RC denote Kullback-Leibler divergence, strategic error rate, and relative completeness. Each performance score is the average of five train-test CV splits. Hyperparameters of the machine learning models are determined by cross-validation within each training data. The number of model parameters (#Params) is the average of the five models for each red and black player.

C.2. Other Performance Measures

Table C9—Prediction Performance Comparison (L1 loss, L2 loss, KL divergence, and strategic error rates)

	Red Players				Black Players			
	L1	L2	KL	SER	L1	L2	KL	SER
Conventional Models								
Constant (Baseline)	1.472	0.854	1.360	0.350	1.466	0.852	1.354	0.562
2-period Nonparametric	1.459	0.851	1.354	0.341	1.451	0.847	1.347	0.539
EWA	1.466	0.851	1.352	0.332	1.464	0.850	1.349	0.559
Serial Correlation								
Order 1	1.467	0.853	1.354	0.350	1.460	0.850	1.349	0.541
Order 4	1.463	0.851	1.351	0.343	1.455	0.848	1.346	0.536
Machine Learning Models								
LASSO	1.454	0.847	1.344	0.324	1.448	0.845	1.339	0.526
Decision Tree	1.459	0.850	1.352	0.343	1.454	0.848	1.348	0.536
LightGBM	1.456	0.848	1.344	0.325	1.452	0.846	1.340	0.527
DNN	1.458	0.849	1.347	0.332	1.455	0.847	1.343	0.534
LSTM	1.437	0.839	1.331	0.308	1.436	0.839	1.329	0.507
Improved Models								
ME1 ($t = 4$)	1.449	0.844	1.336	0.314	1.449	0.844	1.335	0.521
ME2 ($t = 4$)	1.447	0.843	1.335	0.311	1.447	0.843	1.333	0.517
ME2*	1.443	0.842	1.332	0.311	1.445	0.842	1.330	0.516

Notes: Average prediction performance measured in the test data. L1, L2, KL, and SER stand for L1 loss, L2 loss, Kullback-Leibler divergence, and strategic error rates. Each performance score is the average of five train-test CV splits. Hyperparameters of the machine learning models are determined by cross-validation within each training data.

Table C9 lists L1 loss and L2 loss of the selected models listed in Table 8 and 9 measured by test data (averages of the five cross-validation splits). Readers will observe that among all models, the baseline "Constant" model exhibits the poorest performance, while the LSTM model outperforms others across all metrics for both red and black players. Based on this foundation, we define *relative completeness measured by metric ℓ*, RC^{ℓ}, in the following manner:

(C1)

$$(\mathrm{RC}^{\ell} \text{ of a model}) = \frac{\text{(Average loss of the model) - (Average loss of Constant)}}{\text{(Average loss of LSTM) - (Average loss of Constant)}}.$$

In Table C10, we present $\mathrm{RC}^{\mathrm{L1}}$, $\mathrm{RC}^{\mathrm{L2}}$, and $\mathrm{RC}^{\mathrm{SER}}$ along with $\mathrm{RC}^{\mathrm{KL}}$. While

Table C10—Prediction Performance Comparison (Relative Completeness measured by L1 loss, L2 loss, KL divergence, and strategic error rates)

	Red Players				Black Players			
	RC^{L1}	RC^{L2}	RC^{KL}	RC^{SER}	RC^{L1}	RC^{L2}	RC^{KL}	RC^{SER}
Conventional Models								
Constant (Baseline)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2-period Nonparametric	0.372	0.239	0.215	0.210	0.491	0.331	0.297	0.411
EWA	0.177	0.245	0.276	0.429	0.044	0.125	0.195	0.049
Serial Correlation								
Order 1	0.147	0.108	0.204	0.000	0.185	0.133	0.199	0.383
Order 4	0.251	0.209	0.293	0.169	0.345	0.296	0.334	0.464
Machine Learning Models								
LASSO	0.516	0.486	0.562	0.623	0.573	0.522	0.589	0.641
Decision Tree	0.367	0.287	0.273	0.176	0.380	0.281	0.248	0.458
LightGBM	0.448	0.434	0.547	0.594	0.461	0.446	0.556	0.630
DNN	0.388	0.370	0.455	0.430	0.360	0.340	0.432	0.506
LSTM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Improved Models								
ME1 ($t = 4$)	0.661	0.676	0.825	0.850	0.571	0.595	0.753	0.744
ME2 ($t = 4$)	0.713	0.732	0.870	0.928	0.619	0.651	0.848	0.808
ME2*	0.818	0.851	0.961	0.938	0.703	0.750	0.932	0.823

Notes: Relative completeness computed by L1 loss, L2 loss, KL-divergence, and strategic error rates listed in Table C9.

the ranking of model scores is generally consistent across all metrics, some differences exist. For conventional models, the L1 loss metric favors the 2-period Nonparametric model, whereas KL divergence shows a preference for EWA.

ME2* ranks as the second-best performer for both $\mathrm{RC}^{\mathrm{L1}}$ and $\mathrm{RC}^{\mathrm{L2}}$. However, its relative performance sees a significant drop from 96% ($\mathrm{RC}^{\mathrm{KL}}$) to 82-85% for the red player, and from 93% ($\mathrm{RC}^{\mathrm{KL}}$) to 70-75% for the black player.

In contrast, Table C11 displays the Akaike Information Criteria (AIC) and the Bayes Information Criteria (BIC) measured in the entire data. More precisely, we estimate (train) each model using all 2589 pairs in the data and compute the log-likelihood using 26 periods (from period 5 to 30) $\times$ 2589 pairs. Then, the AIC and BIC of a model are calculated as

$$\text{AIC of a model} = -2 \times (\text{log-likelihood}) + 2 \times (\#\text{Params})$$
$$\text{BIC of a model} = -2 \times (\text{log-likelihood}) + (\#\text{Params}) \times \log((\text{sample size})).$$

Both AIC and BIC incorporate a trade-off between the goodness of fit of the model (log-likelihood) and the simplicity of the model (the number of parameters). The model with a lower AIC (BIC) is considered to be preferred.

Table C11 shows that the best model chosen by the AIC (BIC) is quite different compared to those chosen by RC (measured by KL divergence in test data). The machine learning models with large parameter sizes are regarded as bad models according to the two information criteria. One reason is that DNN and LSTM are singular models, which means the Fisher information matrix of a model is a singular matrix. For those models, traditional information criteria are known to be inappropriate in the machine learning literature.

Table C11—Prediction Performance Comparison (AIC, BIC, and RC)

	Red Players				
	#Params	Log-Likelihood	AIC	BIC	RC
Conventional Models					
Constant (Baseline)	3	-18305.8	36617.7	36640.2	0.000
2-period Nonparametric	768	-18223.0	37982.0	43747.9	0.215
EWA	8	-18199.9	36415.8	36475.9	0.276
Serial Correlation					
Order 1	5	-18227.4	36464.7	36502.2	0.204
Order 4	11	-18193.3	36408.6	36491.2	0.293
Machine Learning Models					
LASSO	408.2	-18089.8	36996.1	39296.6	0.562
Decision Tree	40.4	-18201.0	37009.7	39692.4	0.273
LightGBM	30.0	-18095.5	36251.1	36476.3	0.547
DNN	6208.0	-18130.8	48677.6	95285.3	0.455
LSTM	5524.0	-17921.4	46890.9	88364.0	1.000
Improved Models					
ME1	24	-17988.7	36025.4	36205.6	0.825
ME2	38	-17971.2	36018.4	36303.7	0.870
ME2*	126	-17936.5	36125.0	37071.0	0.961
	Black Players				
	#Params	Log-Likelihood	AIC	BIC	RC
Conventional Models					
Constant (Baseline)	0.000	-18231.4	36468.7	36491.3	0.000
2-period Nonparametric	0.297	-18129.2	37794.5	43560.4	0.297
EWA	0.195	-18164.3	36344.6	36404.6	0.195
Serial Correlation					
Order 1	5	-18163.1	36336.1	36373.6	0.199
Order 4	11	-18116.5	36255.1	36337.6	0.334
Machine Learning Models					
LASSO	476.0	-17972.7	35993.4	36173.6	0.589
Decision Tree	41.6	-17940.0	35956.0	36241.3	0.248
LightGBM	30.0	-17911.1	36042.2	36868.0	0.556
DNN	6428.0	-18028.8	36555.3	38423.2	0.432
LSTM	11204.0	-18146.1	36375.3	36687.6	1.000
Improved Models					
ME1	24	-18040.1	36140.2	36365.4	0.753
ME2	38	-18082.7	49021.4	97281.1	0.848
ME2*	110	-17887.6	58183.3	142291.9	0.932

Notes: Log-likelihood shows the sum of the log-likelihood of 26 periods × 2,589 pairs. AIC and BIC are computed by the log-likelihood the number of parameters (#Params). Hyperparameters of the machine learning models are determined by cross-validation within each training data. We repeat the RC (relative completeness) measured by KL divergence of each model shown in Table 8 and 9 for comparison.

D. *Best History Length of Mixing EWA Models*

In this section, we examine the prediction performance of the ME2(t) models across different history lengths t. In addition to using the full training dataset, we report results obtained from training datasets limited to 100, 200, and 400 pairs.

Figure D3 presents the average KL divergence for ME2(t) models with history lengths ranging from $t = 1$ to $t = 20$. The results indicate that the optimal history lengths are $t = 15$ for red players and $t = 13$ for black players. With smaller training datasets, shorter optimal history lengths are identified. The specific optimal values of t are detailed in Table D12.

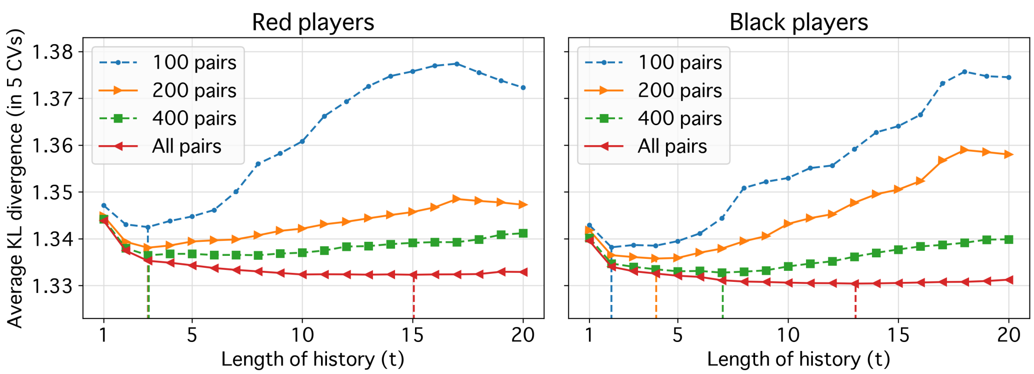


Figure D3. Average KL divergence of ME2(t) models over different t

Note: The average KL divergence of ME2 models over different lengths of history t for various numbers of training data pairs (left: red players, right: black players). Different line styles and colors represent different numbers of pairs in training data: 100, 200, 400, and all pairs. Vertical dotted lines indicate the optimal length of history t for each number of pairs.

Table D12—Best History Length of ME2(t) when Training Data is Small

Number of Pairs in Training Data	Red Players			Black Players		
	Best t	#Params	KL	Best t	#Params	KL
100	3	28	1.342	2	20	1.338
200	3	28	1.338	4	36	1.336
400	3	28	1.336	7	60	1.333
All (2071.2)	15	124	1.332	13	108	1.330

E. Estimated Parameters of EWA Models

E.1. EWA Models

In this section, we provide the complete estimation results of the EWA model, previously omitted in Section III.C. Using the full dataset, we estimate the EWA model via maximum likelihood, along with three variant models: reinforcement learning (RL), belief learning (BL, also known as fictitious play), and nested EWA.

Following Camerer and Ho (1999), we impose the parameter restrictions $\rho = 0$ and $\delta = 0$ in the RL model, and $\rho = \phi$ and $\delta = 1$ in the BL model. To reflect the inherent symmetry among the numbered cards, we also estimate the nested EWA model, whose choice probabilities are defined as:

$$P_i^a(t) \propto \begin{cases} \left(\sum_{k=1,2,3} e^{\frac{\lambda_i}{\eta_i} A_i^k(t-1)} \right)^{\eta_i - 1} \times e^{\frac{\lambda_i}{\eta_i} A_i^a(t-1)} & \text{if } a = 1,2,3 \\ e^{\lambda_i A_i^a(t-1)} & \text{if } a = K. \end{cases}$$

Note that when $\eta_i = 1$, the choice probability simplifies to the standard EWA model form: $P_i^a(t) \propto e^{\lambda_i A_i^a(t-1)}$ for all cards a for all cards a.

Table E13 summarizes the estimation results. The key parameter estimates for the EWA model are $\delta_{\mathrm{R}} = 0.385$ and $\delta_{\mathrm{B}} = 0.000$, indicating that the red player's strategy combines reinforcement learning and belief learning, whereas the black player relies exclusively on reinforcement learning.

We conduct likelihood-ratio (LR) tests to compare the EWA model against RL and BL models, since these two models represent special cases of EWA under the aforementioned parameter restrictions. An LR test is also performed to compare the EWA model with the nested EWA model. The LR test statistics reveal that all differences are statistically significant at the 1% level, although this significance does not directly imply superiority in test-data prediction performance. Indeed, as indicated in Table C8, the standard EWA model for the black player exhibits better predictive performance (lower KL divergence) compared to the nested EWA model.

Table E13—Estimation results of EWA-variant models

Models	Logit			Nested logit
	EWA	RL	BL	Nested EWA
Discount factors				
ϕ_{R}	1.021***	2.527***	1.008***	1.011***
	(0.016)	(0.632)	(0.006)	(0.013)
ϕ_{B}	1.037***	2.314***	0.985***	1.026***
	(0.025)	(0.432)	(0.008)	(0.019)
ρ_{R}	0.928***	$\underline{0.000}$	$\underline{= \phi_{\mathrm{R}}}$	0.912***
	(0.096)			(0.134)
ρ_{B}	1.017***	$\underline{0.000}$	$\underline{= \phi_{\mathrm{B}}}$	0.994***
	(0.005)			(0.004)
Mixing parameters				
δ_{R}	0.385***	$\underline{0.000}$	$\underline{1.000}$	0.356***
	(0.056)			(0.060)
δ_{B}	0.000	$\underline{0.000}$	$\underline{1.000}$	0.000
	(0.071)			(0.057)
Accuracy parameters				
λ_{R}	0.796	0.561***	0.046***	0.899
	(0.742)	(0.151)	(0.004)	(1.106)
λ_B	1.142*	0.656***	0.018***	1.370**
	(0.601)	(0.120)	(0.005)	(0.651)
Nest parameters				
η_{R}	$\underline{1.000}$	$\underline{1.000}$	$\underline{1.000}$	1.410***
				(0.247)
η_{B}	$\underline{1.000}$	$\underline{1.000}$	$\underline{1.000}$	1.330***
				(0.247)
Initial values				
$A^1_{\mathrm{R}}(0)$	−0.518	−1.059**	−7.477***	−0.919
	(0.555)	(0.428)	(0.806)	(1.182)
$A^2_{\mathrm{R}}(0)$	−0.677	−1.350**	−9.963***	−1.105
	(0.723)	(0.541)	(0.942)	(1.418)
$A^3_{\mathrm{R}}(0)$	−0.777	−1.552**	−11.447***	−1.222
	(0.823)	(0.611)	(1.033)	(1.562)
$A^1_{\mathrm{B}}(0)$	−0.484	−1.048***	−20.149***	−0.674*
	(0.323)	(0.312)	(5.190)	(0.389)
$A^2_{\mathrm{B}}(0)$	−0.644	−1.398***	−27.542***	−0.840*
	(0.427)	(0.420)	(7.084)	(0.477)
$A^3_{\mathrm{B}}(0)$	−0.670	−1.447***	−28.275***	−0.866*
	(0.446)	(0.436)	(7.075)	(0.492)
$N_{\mathrm{R}}(0)$	7.071	$\underline{1.000}$	$\underline{1.000}$	7.917
	(7.144)			(9.563)
$N_{\mathrm{B}}(0)$	6.692	$\underline{1.000}$	$\underline{1.000}$	8.304*
	(4.167)			(4.475)
Number of Observations	77670	77670	77670	77670
Number of Parameters	8	5	5	9
Log Likelihood (Red)	-105059.3	-105569.0	-105160.5	-105010.2
Log Likelihood (Black)	-104831.7	-104945.9	-105319.1	-104804.5
LR Statistic (Red)		1019.4***	202.4***	98.3***
LR Statistic (Black)		228.3***	974.7***	54.5***

Notes: Maximum likelihood estimates for parameters of the EWA model variants. Standard errors are reported in parentheses. Underlined values indicate parameters determined by model restrictions rather than estimation. The number of observations and parameters refer to each of the models for red and black players, respectively. The LR Statistic represents the likelihood ratio test statistic comparing each variant model against the original EWA model. *p<0.1; **p<0.05; ***p<0.01.

E.2. Mixing EWA Models

We report the estimated parameters of the improved models for ME2($t = 4$) (Table E14) and ME2* (Table E15). Although most additional parameters are insignificant even at the 10% level, both ME2($t = 4$) and ME2* outperform the original EWA in predicting the test data.

TABLE E14—ESTIMATED PARAMETERS OF ME2 ($t = 4$)

(a) Red Players

EWA parameters	δ_R	ρ_R	ϕ_R	λ_R	$A^1_R(0)$	$A^2_R(0)$	$A^3_R(0)$	$N_R(0)$
	0.230	0.882***	1.004***	0.844***	−0.448***	−0.616	−0.718	6.284*
	(3.721)	(0.048)	(0.003)	(0.076)	(0.105)	(0.509)	(0.521)	(3.294)
Additional parameters	$a = K$				$a = 1, 2, 3$			
	$\alpha^{a\tau}_{R1}$	$\alpha^{a\tau}_{R2}$	$\gamma^{a\tau}_{R1}$	$\gamma^{a\tau}_{R2}$	$\alpha^{a\tau}_{R1}$	$\alpha^{a\tau}_{R2}$	$\gamma^{a\tau}_{R1}$	$\gamma^{a\tau}_{R2}$
$\tau = 1$	0.075	0.096	−0.177		−0.190	0.082	−0.015	
	(1.894)	(1.724)	(0.526)		(2.019)	(1.804)	(0.112)	
2	0.123	0.222	0.014	0.133	0.257	0.220	0.042	−0.020
	(0.512)	(0.760)	(0.451)	(0.516)	(0.340)	(0.217)	(0.176)	(0.130)
3	0.290	0.148	0.149	0.004	0.508	0.121	0.006	−0.010
	(1.220)	(0.530)	(0.993)	(0.609)	(2.792)	(0.385)	(0.278)	(0.462)
4	0.360	−0.077	0.170	−0.129	0.330	−0.009	−0.076	−0.050
	(4.890)	(1.492)	(8.365)	(1.065)	(40.514)	(0.660)	(5.309)	(0.806)

(b) Black Players

EWA parameters	δ_B	ρ_B	ϕ_B	λ_B	$A^1_B(0)$	$A^2_B(0)$	$A^3_B(0)$	$N_B(0)$
	0.000	0.933***	0.989***	1.064***	−0.360***	−0.591***	−0.635***	5.177***
	(0.108)	(0.005)	(0.009)	(0.044)	(0.132)	(0.073)	(0.058)	(0.344)
Additional parameters	$a = K$				$a = 1, 2, 3$			
	$\alpha^{a\tau}_{B1}$	$\alpha^{a\tau}_{B2}$	$\gamma^{a\tau}_{B1}$	$\gamma^{a\tau}_{B2}$	$\alpha^{a\tau}_{B1}$	$\alpha^{a\tau}_{B2}$	$\gamma^{a\tau}_{B1}$	$\gamma^{a\tau}_{B2}$
$\tau = 1$	0.076	0.194	−0.098		−0.256	0.017	−0.111*	
	(1.058)	(1.042)	(0.237)		(1.028)	(1.037)	(0.061)	
2	0.156	0.197	0.038	−0.127	0.244*	0.241*	−0.038	0.115
	(0.265)	(0.238)	(0.231)	(0.296)	(0.131)	(0.139)	(0.068)	(0.132)
3	0.187	0.030	0.038	−0.045	0.265	0.139	−0.041	−0.006
	(0.375)	(0.314)	(0.289)	(0.499)	(1.178)	(0.427)	(0.102)	(0.235)
4	0.265	−0.113	−0.075	0.038	0.964	−0.014	−0.011	0.040
	(3.626)	(0.360)	(0.332)	(1.617)	(2.658)	(0.392)	(0.124)	(0.474)

Notes: Maximum likelihood estimates of the parameters of mixing EWA models. Robust standard errors are in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table E15—Estimated Parameters of ME2* (Red Players)

(a) Red Players

EWA parameters	δ_{R}	ρ_{R}	ϕ_{R}	λ_{R}	$A^1_{\mathrm{R}}(0)$	$A^2_{\mathrm{R}}(0)$	$A^3_{\mathrm{R}}(0)$	$N_{\mathrm{R}}(0)$
	0.310	0.895***	1.007***	0.814***	−0.496	−0.669***	−0.774***	6.646***
	(1.000)	(0.008)	(0.007)	(0.030)	(0.402)	(0.168)	(0.138)	(0.373)
Additional parameters	$a = K$				$a = 1, 2, 3$			
	$\alpha^{a\tau}_{\mathrm{R1}}$	$\alpha^{a\tau}_{\mathrm{R2}}$	$\gamma^{a\tau}_{\mathrm{R1}}$	$\gamma^{a\tau}_{\mathrm{R2}}$	$\alpha^{a\tau}_{\mathrm{R1}}$	$\alpha^{a\tau}_{\mathrm{R2}}$	$\gamma^{a\tau}_{\mathrm{R1}}$	$\gamma^{a\tau}_{\mathrm{R2}}$
$\tau = 1$	0.069	0.090	−0.176		−0.179	0.091	−0.016	
	(2.084)	(2.187)	(0.360)		(1.961)	(2.050)	(0.139)	
2	0.128	0.222	0.016	0.134	0.262	0.218	0.042	−0.019
	(0.469)	(0.673)	(0.715)	(0.354)	(1.023)	(0.220)	(0.199)	(0.148)
3	0.289	0.146	0.152	0.004	0.510	0.118	0.007	−0.010
	(0.937)	(0.585)	(0.882)	(0.720)	(9.222)	(0.706)	(0.288)	(0.176)
4	−0.002	−0.019	0.175	−0.075	0.074	0.040	−0.006	−0.090
	(1.404)	(0.977)	(4.154)	(0.759)	(14.689)	(0.855)	(1.428)	(0.177)
5	0.004	0.017	0.090	−0.014	0.220	−0.036	−0.058	0.073
	(9.353)	(0.733)	(15.356)	(1.117)	(7.501)	(0.747)	(3.698)	(0.335)
6	0.462	−0.055	−0.343	−0.082	0.047	0.009	−0.118	−0.061
	(9.136)	(1.599)	(14.391)	(4.706)	(12.692)	(0.822)	(7.987)	(0.448)
7	0.051	−0.060	0.120	0.119	1.029	−0.053	−0.021	0.008
	(2.804)	(1.894)	(4.196)	(5.111)	(24.607)	(0.538)	(4.727)	(0.626)
8	0.306	−0.140	0.869	−0.209	0.634	0.072	0.349	0.236
	(1.748)	(7.754)	(3.877)	(28.005)	(15.880)	(0.852)	(13.629)	(0.700)
9	0.171	0.178	−0.520	−0.057	−0.638	−0.094	−0.023	−0.168
	(9.849)	(2.423)	(9.037)	(31.503)	(10.890)	(0.663)	(38.267)	(0.685)
10	1.385	−0.311	−0.038	0.180	−0.037	−0.066	−0.495	−0.159
	(5.519)	(41.644)	(8.879)	(104.568)	(20.359)	(0.854)	(28.645)	(0.811)
11	0.722	−0.170	−0.890	−0.071	−0.008	−0.037	0.034	−0.214
	(14.586)	(25.781)	(103.658)	(61.595)	(8.354)	(0.508)	(8.544)	(0.794)
12	−0.480	0.296	0.815	−0.433	1.860	0.053	−0.219	0.519
	(6.155)	(51.490)	(102.494)	(48.341)	(5.815)	(1.271)	(8.400)	(2.703)
13	1.576	−0.434	−0.180	0.153	1.561	−0.068	−1.143	−0.288
	(1.556)	(53.412)	(8.205)	(5.708)	(18.143)	(1.511)	(3.055)	(5.364)
14	1.263	0.240	−0.892	−0.135	1.417	−0.098	0.084	−0.451
	(5.159)	(53.211)	(68.606)	(4.209)	(29.540)	(2.902)	(6.325)	(3.378)
15	1.172	−0.581	1.163	0.471	1.398	−0.079	−1.684	0.977
	(4.589)	(28.655)	(74.036)	(10.713)	(17.968)	(0.759)	(2.454)	(3.557)

Notes: *p<0.1; **p<0.05; ***p<0.01.

(b) Black Players

EWA parameters	δ_{R}	ρ_{R}	ϕ_{R}	λ_{R}	$A^1_{\mathrm{R}}(0)$	$A^2_{\mathrm{R}}(0)$	$A^3_{\mathrm{R}}(0)$	$N_{\mathrm{R}}(0)$
	0.000	0.970***	1.006***	1.099***	−0.334***	−0.562***	−0.608***	5.199***
	(0.076)	(0.004)	(0.009)	(0.046)	(0.110)	(0.069)	(0.052)	(0.289)
Additional parameters	$a = K$				$a = 1, 2, 3$			
	$\alpha^{a\tau}_{\mathrm{R1}}$	$\alpha^{a\tau}_{\mathrm{R2}}$	$\gamma^{a\tau}_{\mathrm{R1}}$	$\gamma^{a\tau}_{\mathrm{R2}}$	$\alpha^{a\tau}_{\mathrm{R1}}$	$\alpha^{a\tau}_{\mathrm{R2}}$	$\gamma^{a\tau}_{\mathrm{R1}}$	$\gamma^{a\tau}_{\mathrm{R2}}$
$\tau = 1$	0.097	0.211	−0.098		−0.264	0.008	−0.112	
	(0.676)	(0.552)	(0.243)		(0.839)	(0.764)	(0.153)	
2	0.167	0.193	0.043	−0.135	0.250*	0.238*	−0.039	0.113
	(0.285)	(0.292)	(0.333)	(0.486)	(0.134)	(0.141)	(0.150)	(0.153)
3	0.193	0.023	0.043	−0.057	0.269	0.135	−0.040	−0.008
	(0.446)	(0.337)	(0.388)	(0.561)	(1.360)	(0.445)	(0.175)	(0.330)
4	0.133	0.003	−0.033	0.074	0.266	−0.005	−0.015	0.026
	(0.807)	(0.607)	(0.741)	(0.814)	(2.786)	(0.516)	(0.407)	(0.580)
5	0.105	−0.047	0.046	−0.183	0.834	0.029	−0.077	0.029
	(2.090)	(0.793)	(0.676)	(1.228)	(25.224)	(0.490)	(0.250)	(0.895)
6	−0.048	−0.062	−0.070	−0.123	−0.527	0.032	0.063	−0.055
	(3.631)	(0.930)	(0.688)	(8.002)	(100.963)	(1.062)	(0.476)	(6.132)
7	0.467	−0.042	0.026	0.200	1.502	−0.039	0.164	0.096
	(3.728)	(2.312)	(1.019)	(4.873)	(115.510)	(0.874)	(1.346)	(6.806)
8	−0.313	−0.147	−0.062	−0.102	0.333	−0.018	−0.077	−0.253
	(6.476)	(1.581)	(1.118)	(39.659)	(2.632)	(0.970)	(1.713)	(10.644)
9	0.068	−0.213	−0.014	0.131	−0.836	−0.032	−0.029	−0.183
	(7.308)	(7.371)	(2.119)	(34.732)	(9.611)	(0.542)	(1.125)	(2.017)
10	0.967	0.131	−0.015	0.216	1.055	−0.042	0.016	−0.284
	(4.176)	(13.010)	(2.528)	(27.977)	(3.996)	(0.439)	(1.708)	(4.036)
11	0.055	−0.543	−0.009	0.021	1.434	−0.005	−0.067	0.527
	(0.973)	(4.738)	(3.671)	(10.290)	(2.113)	(0.869)	(5.612)	(7.747)
12	0.451	0.475	−0.106	0.157	1.214	−0.009	0.028	0.467
	(0.751)	(34.726)	(31.318)	(6.021)	(20.709)	(0.701)	(4.552)	(46.812)
13	−0.566	−0.552	−0.320	0.466	1.130	−0.203	0.347	−1.202
	(1.694)	(36.281)	(33.586)	(48.211)	(6.072)	(0.375)	(31.985)	(39.794)

Notes: *p<0.1; **p<0.05; ***p<0.01.