**Learning in Games: Evidence from the 2014 Pokerstars Price Increase**

**Abstract**

*We exploit a natural experiment in online poker to identify the empirical credibility of models of learning in game theory to predict player behavior. Controlling for player heterogeneity, we find that player experience based learning rules (such as reinforcement learning or regret minimization) do poorly to explain human learning in complex games with noisy payoffs.*

1. **Introduction**

On October 24th, 2014, Pokerstars, the largest online poker room[[1]](#footnote-1), announced a large increase in the effective price to play poker to take effect 10 days later. This price increase was exogenous from the poker players’ perspective[[2]](#footnote-2) and reversed approximately two months later for reasons unspecified by Pokerstars management.

This price shock serves as a natural experiment to examine the behavioral response of poker players. Because the cost to play online poker is implemented directly on the money wagered by poker players and is collected unevenly through the poker game tree, the price change affects the expected value of some strategies and not others. This serves as a rare instance of a natural field experiment in behavioral game theory.

We find that, while there is a significant exogenous change in the expected value of some decisions relative to others, poker players under-adapt their strategies, at a significant monetary cost. Players however do adapt their strategies significantly depending on the type of player they are up against. This finding implies that players do not use payoff information significantly in learning when the payoff deviations are not salient, even when the deviation is large.

There is a tendency towards studying simpler games in empirical game theory. This is usually done to ensure all studied players understand the rules of the game and to keep the analysis tractable. No-limit Texas Hold’em, the game played in our dataset, stands in sharp contrast to this trend with a state space of more than 10140 possible situations (Johansen, 2013).

This complexity makes many usual avenues of analysis impossible – simply indexing possible strategies can be intractable However, full scale poker games like the one we are studying have historically been considered something of a microcosm of human strategic behavior under imperfect information. For instance, Von Neumann, arguably the founder of game theory, once stated, referring to poker: *“Real life consists of bluffing, of little tactics of deception, of asking yourself what is the other man going to think I mean to do. And that is what games are about in my theory.”* (Bonowski, 1973).

The paper is organized as follows. **Section 2** reviews the literature of learning dynamics in games. In **section 3**, we discuss the data and the 2014 price change. In **section 4**, we discuss unobserved heterogeneity in poker players and methods to deal with the problem. In **section 5**, we discuss the game structure of poker and the effect of the price change on said structure. In **section 6.1**, we present empirical results in the change in value of decisions for players from the price change. In **section 6.2,** we present empirical results on the change in behavior from players. **Section 7** estimates how large the strategy change should have been given empirically observed values. **Section 8** concludes.

1. **Learning in games**

Assuming players are at equilibrium in a game (Nash or otherwise) is normally a very strong assumption. This assumption is instead a completely unreasonable one in our case, as the equilibrium is cannot be computed by the best current supercomputers, let alone human players.

Instead, assuming players are playing in the process of learning their strategies is a more reasonable approach. However, Fudenberg & Levine (2016) note that the predictive power of learning models depends on circumstances, and that we do not yet have a good understanding of which performs well when.

The problem of circumstantial performance of learning models is formalized by strong negative theoretical results. Hart & Mas-Collel (2003) show that no strategy adjustment rules that consider only one’s own payoff function can lead to Nash equilibrium in all games. Furthermore, Babichenko & Rubinstein (2016) prove that no learning rule using information from previous games will converge to an -Nash equilibrium in reasonable timeframes for every possible game.

There are two classes of learning models in economics. The first is experience based learning, where a player learns based on his experience playing the game. The second is observation based learning, where a player learns the game by observing the behavior of other players.

The two most popular models of learning use to predict player behavior in repeated games are belief learning and reinforcement learning. These represent observation and experience based learning respectively. In belief learning**,** a player learns about what his opponent is likely to do and uses this belief to decide on his strategy. The most popular belief learning model is Fictitious Play (Shapley, 1964), where players best respond to the empirical mixed strategy observed by the opponent in previous play.

Reinforcement learning (Roth & Erev 1995) is inspired by the psychology literature and has a player update his strategy based on the empirical utility the strategy provided. In the simple reinforcement learning model, the only action whose mixed strategy probably is updated in a round of a repeated game is the one for the action taken in that round.

Belief learning and reinforcement learning are combined by Camerer & al. (1999) in a model called “Experience Weighed Attraction” (EWA) which can act as either model or a combination of both based on a free parameter. EWA models transcend “observation vs experience” classification and have had success in predicting players behavior in simple games (SOURCE). However, fitting a EWA model to full scale poker in our dataset is computationally intractable. Moreover, EWA’s reinforcement learning side relies on the simple reinforcement learning model (SOURCE), which is unreasonable in large games.

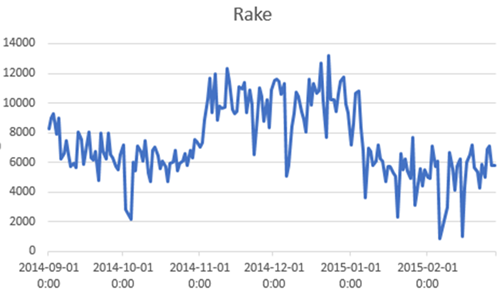
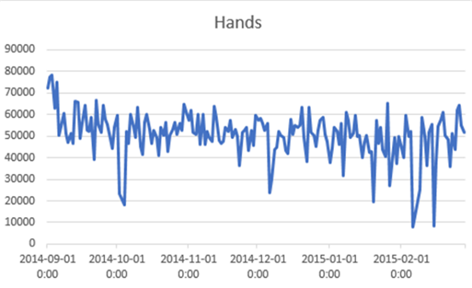
The critical problem with simple reinforcement learning in games with a large action space is that actions unexplored will never be reinforced. For example, Mohlin, Ostling and Wang (2015) use this argument to demonstrate that simple reinforcement learning cannot explain empirical behavior in field data in a comparatively small game. In the game studied here, where it’s impossible to explore the action space in a human lifespan, the model is clearly inadequate

Another popular learning rule is called *regret matching*, introduced by Hart & Mas-Collel (2000), has been outstandingly effective to create computer poker agents. In regret matching, a player plays a strategy proportional to the counterfactual accumulated payoff had he played his other possible strategies in past rounds. Intuitively, instead of myopically considering the action taken, the regret matching player can reflect ex-post of the counterfactual performance of his other options. The regret matching algorithm has been used to solve the Nash equilibrium Limit-Hold’em (UALBERTA 2014), one of the largest games ever solved. It is also used in artificial game playing intelligence to beat the top human experts in 1 vs 1 No-Limit Texas Hold’em (Brown and Sandholm, 2017).

Minimizing regret has many desirable properties for a learning rule. First, strategies that minimize regret have proven performance guarantees: those strategies’ value exceeds the minimax value of the game in all zero-sum games (Blum & Mansour, 2007). Moreover, regret matching, like fictitious play, are part of the general class of *multiplicative weights update* (MWU) algorithms (SOURCE) which have proven performance guarantees even against adversaries who know the player’s strategy and update method. MWU algorithms are also proven to converge to Nash equilibria in all zero sum games. For these reasons, regret is considered a basic notion in the literature in prediction and learning in games (Cesa-Bianchi & Lugosi, 2006) and is used to model off equilibrium behavior in various games, for example in online auctions (Caragiannis et al., 2015). However, regret matching has not been as popular as EWA in empirically predicting the behavior of players. This is likely because it makes very sharp predictions of future play which turn out to poorly forecast human player behavior.

While we will not test specific models of learning, we can use our data to coarsely test the empirical performance of experience-based learning rules against observation-based rules by exploiting the exogenous change in payoffs.

1. **Data and the 2014 Pokerstars price change**

The data is collected by an poker datamining service, *hhdealer.com*. The data details 4.9 million hands played by 12,000 players, between 09/01/2014 and 02/28/2015 on Pokerstars in 1 v. 1 Texas Hold’em tables at two different stakes ranging in buy-in from $80 to $400.

**Figure 1**. *Daily money collected from players and daily poker hands played in the data. Increase in price is visible between 11/03/2014 and 01/06/2015.*

These games were chosen because they are some of the lowest stakes where professionals online poker players are known to play full time. As such, the price change should have the largest relative effect at these tables. 1 vs 1 tables were chosen because strategic differences are more easily identifiable in this format.

The data is collected in raw text format, logged by Pokerstars for archival purposes. This raw text is parsed and processed into numerical form by a poker database software called *PokerTracker*. Importantly, the data maintains information on all strategic decisions taken by the players in our population. It observes the frequency at which strategic actions are taken, which represents the mixed strategy of a player against his opponent population.[[3]](#footnote-3)

The price increase was introduced by Pokerstars on *November 3rd, 2014* and repealed on *January 6th, 2015.* **Figure 1** shows the aggregate number of hands played each day and the amount collectively lost by players (conversely, the amount made by Pokerstars) in our data’s population, where the price increase is visible. There is no clearly visible change in the amount of hands played in our population.

The initial price increase should be considered exogenous from the players point of view. However, the repeal of the price increase was only explained by Pokerstars as being based on *“additional analysis and consideration” [[4]](#footnote-4)* and should be considered an endogenous response to the effects of the price change. This is initially puzzling, given that the price increase has no measurable impact on the amount of hands of poker played or minutes at the table (see figure 1 for example). This conundrum is resolved with further analysis which shows that lower skill players overwhelmingly shoulder the burden of the price increase (see section 4 and section 6.1)

1. **Heterogeneity in poker players**

Like other popular table games, poker has a population of highly skilled professional players. Professional poker players are also reputed for having desirable behavioral traits. For instance, Levitt, List & Reiley (2010) use professional players as a test group with “unrivaled experience applying analytical thought to card games.” Linnet et al. (2010) find experienced poker players have lower cognitive bias in probability estimation and choice making than unexperienced players.

In game playing artificial intelligence, Moravcik et al. (2017) as well as Brown and Sandholm (2017) use professional human players as the ultimate benchmark against which to test their computer poker agents.

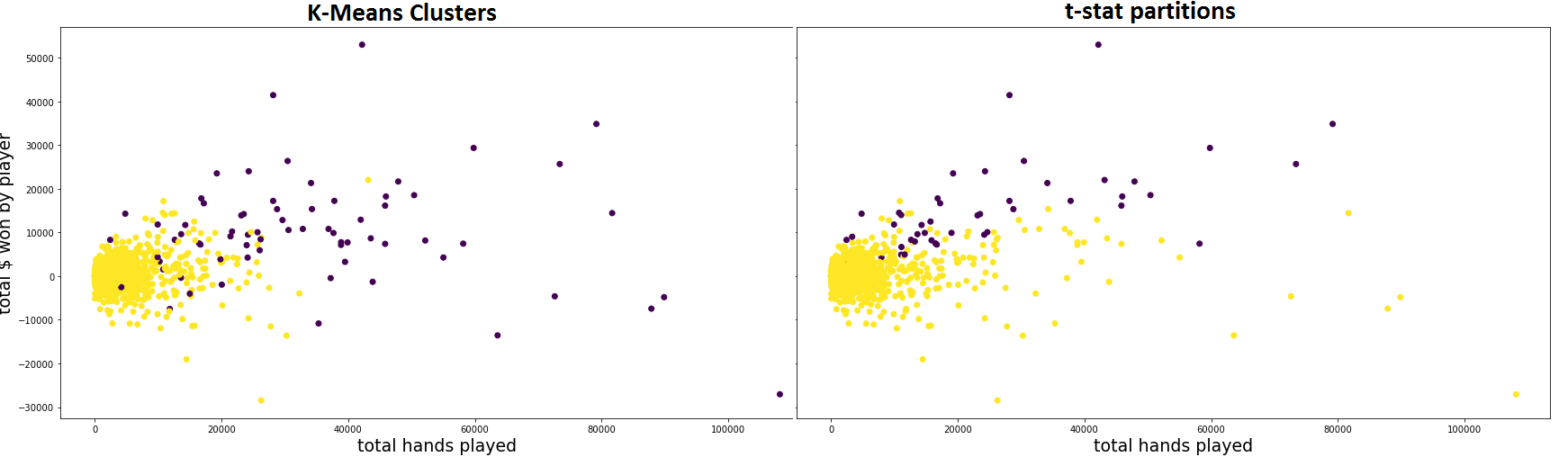
Because professional players present in our dataset are qualitatively different and *a priori* unobservable, the heterogeneity of players is dealt with in the following ways

First, of the 12,000 players in our population, approximately 2500 player played both during the price increase and either before or after. These 2500 players are labeled our **treatment group**. The 9,500 remaining players only played in one of the three treatment periods and are labeled as the non-treatment group, or **ephemeral players**.

Three alternative methods are used to control for the unobserved heterogeneity in the treatment group:

The first method is to *cluster* the dataset into **K** groups, by using a clustering algorithm. Variants of this method have recently gained popularity in econometrics to control for unobserved heterogeneity, for example in Lin & Ng (2012), Bonhomme & Manresa (2015) and Bonhomme, Lamandon & Marsenna (2017). Econometricians generally favor the K-means algorithm (Steinley, 2006) to achieve the clustering. We also use the K-Hierarchical agglomerative clustering algorithm (Hastie & Tishranie, 2009) as an alternative cluster specification.

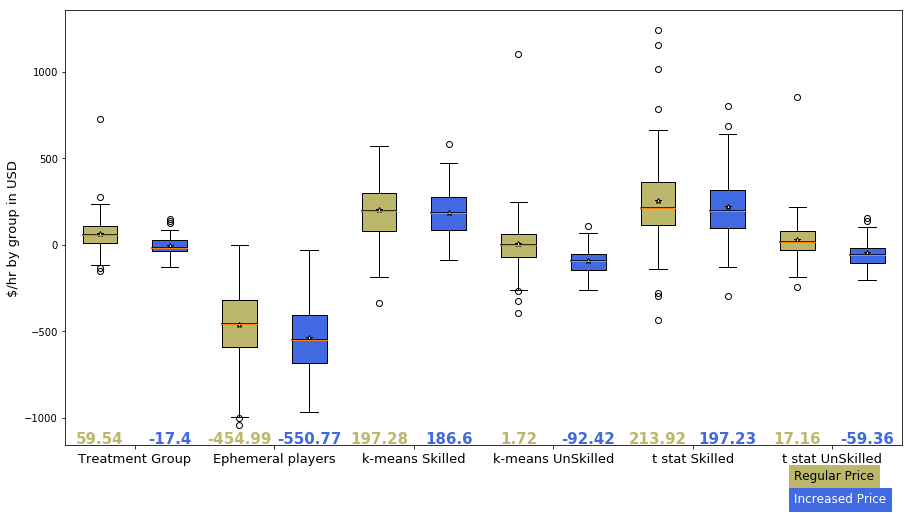
The clusters presented for the rest of the paper use 15 variables of interest to cluster individuals: The total number of hands played, minutes played, and total winnings, the average of the previous three variables per session played, as well as 9 statistics of strategic behavior at the table pulled from *PokerTracker[[5]](#footnote-5)*. These are chosen to best represent player behavior. Results using other cluster specifications are shown in the appendix.



**Figure 2.** *Clusters of high skill players (purple) and others (yellow) in the treatment group, as defined by the k-means algorithm and likelihood of being a winning player, respectively*

The second method is to employ a latent class model by using a finite mixture of Gaussian distributions (SOURCE). This method retrieves the parameters for each group, but leaves it difficult to retrieve group membership for individuals in the dataset. Because the behavior of individuals in the dataset is of interest to us, this method will be used mainly as a robustness check.

The last method is to partition the dataset between players who are likely professional players by a t-statistic on their likelihood of being winning players. Because players who played very few hands may pass this test we impose the additional restriction that the players labeled “professionals” by this method have played more than 60 hours in the 6 months observed in the data (a generous average of 2.5 hours per week). The t-statistic is calculated by:

Where *wintot* is the total poker winnings (positive or negative) of the player in the dataset in US dollars, *handstot* is the total number of hands played and *SE* is the standard deviation in US dollars of winnings per hand. A player is considered a professional by this method if he has played enough hours and has *t > 1.645*.

**Figure 3.** Average $ earned per hour playing poker per group during regular and increased price periods. High skill players seem unaffected, unlike low skill players.

**Figure 2** plots the hands played and winnings in US dollars for the treatment group in our dataset. On the left,purple dots are individuals classed in the high skill cluster by the K-means algorithm used in the 15 dimensions described above projected into two dimensions. On the right, purple players are professional as defined by the statistical test above.

Next, **figure 3** demonstrates the change in players’ hourly wage by group. High skill players are unaffected by the change in price, while low skill players in the treatment group and ephemeral players are heavily affected. These stylized facts are robust to specifications of the clustering method and restrictions on the number hours played for statistical partitioning.

Low skill players in the treatment group are especially affected by the change, going from marginal winners as a group to clear losers, losing upwards of $50 per hour at the table under the increased price. Ephemeral players as a group are clear losers both before and after the change, losing hundreds of dollars per hour at the table.

These results explain the decision by Pokerstars to reverse the price change despite its profitability. The net depositors on the site (low skill and ephemeral players) are shouldering the entire burden of the price change, while net creditors (professional players) are unaffected.

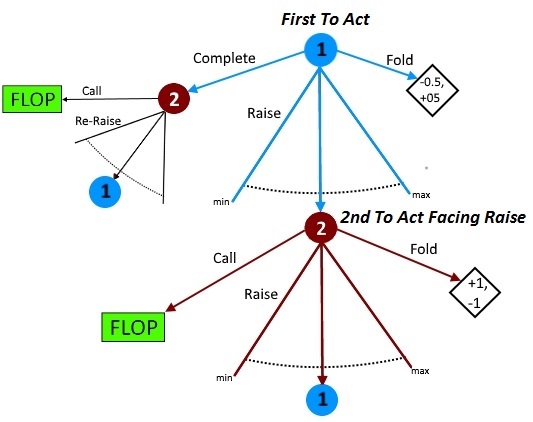
The reason why high skill players’ bottom line is unaffected by the price change is discussed in section 6.2.

1. **The effect of the price change on poker’s game structure**

Most online poker rooms, including Pokerstars, charge poker players using a system called “the rake”, where small amount of money removed from the contested pot of money at each hand and appropriated as revenue for the host of the game. The rake is collected at a rate of 5% of the size of the pot up to a maximum of $0.5 per hand in our sampled games. This maximum is increased to $1 during the period of increased pricing.

Pokerstars employs a policy called “*No Flop, No Drop”*, which states that hands that conclude before a flop is dealt are not charged. A flop is dealt when the first round of betting concludes without a player forfeiting by electing to fold his private cards.

For a constant sum game, any subtree of the game will have the same strategies be optimal regardless of the constant, if the relative size of payoffs is respected. This is the case for the change in price after the flop in our data: if two players reach a post-flop situation with the same information and possible cards held before or after the price change, their strategies should not change.

Because of the intractable complexity of the game, we will concentrate on a restricted set of the most commonly encountered decisions.

**Figure 4.** *Simplified game tree of the first decisions in 1 vs 1 Texas Hold’em poker, after players have been dealt their private cards.*

**Figure 4.** shows the simplified game tree of the first few decisions after receiving cards. End nodes labeled “flop” lead to subtrees which have strictly worse payoffs for both players under the higher price. After player 1’s first decision node, all further decision nodes share similar structure to the “facing raise” node. Choosing to fold leads to the same payoff regardless of Pokerstars’ current price, call leads to a flop node, and raise leads the opposing player to a similar situation.

The node we are interested in is labeled *“2nd to act facing raise.”* It is the node where the choice between strategies has the largest change in expected value from the price increase implemented at the flop. Moreover, it is the most common node to be reached barring the initial decision node. Electing to “complete” when first to act is rarely seen in the data as it is considered a bad strategy by poker players[[6]](#footnote-6), so most poker hands begin with a raise by player 1 and player 2 facing the raise.

* 1. **Empirical effect on expected value of strategies**

Even if the expected value of continuing the hand when 2nd player to act facing a raise should diminish in theory, we would prefer empirical confirmation. To verify this, we regress the average winnings when a player is in this situation in a panel model with fixed effects:

(1)

This regression is run separately on partitions of the dataset as discussed in section 3. is the average value of a re-raise or call action when second to act for player *i* at time *t.* Each observation is a continuous poker “session” starting at the first hand and stopping when one of the two players exits the virtual poker table. Fixed effect controlling for player characteristics are and controls for the opponent are . The variable indicating the increase in price is *Treat* and controls for the skill group of the opponent, as well as interaction effects between the two are found in *X.* Seasonality controls based on time (time of day, day of the week, etc.) were found insignificant in all specifications.

**Table 1** shows the average winnings of each group of players (as defined in section 3) when electing to call or re-raise in this position during both pricing periods. The expected value is denominated in the minimum bet of the game to control for the heterogenous stakes in the games in dollar terms (the minimum bet is either $2 or $4 in our data).

We can see that there is a large effect on low-skill regular players, and no significant effect on high skill players. Ephemeral players are also unaffected by the change, even when aggregated as a group. Importantly, low skill players’ expected value goes from slightly positive to significantly negative, implying that many individual holdings become unprofitable to continue playing with.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***BB***  ***continue***  ***EV*** |  | **Constant** | **Increased**  **Price** | **High Skill Opponent** | **Ephemeral Opponent** | **Ephemeral**  **Opponent**  **x Treat** | **Low skill**  **x Treat** | **High Skill x treat** |
| **Treatment**  **Group** | **All** | 0.555\*\*\*  (0.151) | -0.582\*\*  (0.238) | -1.410\*\*\*  (0.285) | -0.017  (0.190) | 0.468  (0.465) | -0.271  (0.319) | -0.311  (0.427) |
| **High skill**  **(k-means)** | 2.291\*\*\*  (0.321) | -0.325  (0.371) | - 1.567\*\*\*  (0.486) | -1.007\*\*  (0.363) | 0.741  (0.838) | -0.630  (0.573) | 0.304  (0.600) |
| **Low skill**  **(k-means)** | 0.059  (0.171) | -0.711\*\*  (0.321) | -1.861\*\*\*  (0.362) | 0.020  (0.230) | 0.093  (0.611) | 0.026  (0.407) | -0.736  (0.592) |
| **High skill**  **(t-stat)** | 2.531\*\*\*  (0.426) | 0.503  (0.549) | -0.688  (0.669) | -1.010\*\*  (0.478) | -2.053\*  (1.090) | 1.378\*  (0.785) | -0.875  (0.943) |
| **Low skill**  **(t-stat)** | 0.287\*  (0.162) | -0.653\*\*  (0.267) | -1.675\*\*\*  (0.316) | 0.003  (0.210) | 0.474  (0.529) | -0.270  (0.355) | -0.383  (0.480) |
| **Ephemeral**  **Players** |  | -3.492\*\*\*  (0.318) | 0.000  (0.000) | 0.369  (0.564) | 0.642  (0.391) | 0.000  (0.000) | 0.000  (0.000) | 0.000  (0.000) |

**Table 1.** *Expected value of electing to call or raise when second to act facing a raise. Standard errors in parentheses. Rows are estimates for different groups of players. Columns are the effect of different opponent types and price changes on expected value.*

**6.2 Strategic behavior**

We now turn attention to players’ strategic behavior. **Table 2** presents results using the same fixed effects panel specification as model (1) in section 6.1, with average frequency at which a player elects to fold facing a raise when second to act as dependant variable. This position is chosen for the reasons discussed in section 5.

The average frequency is in the *constant* column, hovering around 70% for all player types. This implies players elect to call or re-raise when second to act about 30% of the time when facing a low skill player in the treatment group.

All players elect to fold significantly more often when facing a player from the ephemeral group. The difference in skill of ephemeral players is apparently evident to other players, even other ephemeral players. All player types, except ephemeral players, also play somewhat differently when facing a high skill than a low skill opponent from the treatment group.

The change in price has no effect on players’ strategies when second to act facing a raise. This is also true for players first to act (no change in raise or fold frequencies) as well as for any other strategy tested anywhere in the game tree. Tables showing similar results for other nodes in the game tree are in the

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Fold BB %*** |  | **Constant** | **Increased**  **Price** | **High Skill Opponent** | **Ephemeral Opponent** | **Ephemeral Opponent x treat** | **Low Skill Opponent x treat** | **High Skill x treat** |
| **Treatment**  **Group** | **All** | 69.271\*\*\*  (0.122) | 0.174  (0.174) | 1.951\*\*\*  (0.221) | 11.410\*\*\*  (0.164) | 0.186  (0.395) | 0.261  (0.244) | -0.086  (0.304) |
| **High skill**  **(k-means)** | 72.732\*\*\*  (0.234) | 0.412  (0.291) | 1.334\*\*\*  (0.357) | 13.172\*\*\*  (0.280) | -0.122  (0.692) | 0.167  (0.455) | 0.250  (0.465) |
| **Low skill**  **(k-means)** | 68.282\*\*\*  (0.141) | 0.106  (0.226) | 1.235\*\*\*  (0.282) | 9.174\*\*\*  (0.205) | 0.236  (0.521) | 0.206  (0.302) | -0.100  (0.406) |
| **High skill**  **(t-stat)** | 72.953\*\*\*  (0.234) | -0.059  (0.456) | 2.153\*\*\*  (0.532) | 13.802\*\*\*  (0.393) | -0.500  (1.039) | -0.561  (0.707) | 0.502  (0.734) |
| **Low skill**  **(t-stat)** | 68.771\*\*\*  (0.131) | 0.193  (0.192) | 1.683\*\*\*  (0.242) | 10.374\*\*\*  (0.182) | 0.565  (0.446) | 0.359  (0.266) | -0.166  (0.338) |
| **Ephemeral**  **Players** |  | 71.721\*\*\*  (0.223) | 0.000  (0.000) | -0.269  (0.441) | 16.010\*\*\*  (0.279) | 0.000  (0.000) | 0.000  (0.000) | 0.000  (0.000) |

**Table 2.** *Strategies of players against other player types and during price change. Standard errors in parentheses. Rows are estimates for different groups of players. Columns are the effect of different opponent types and price changes on mixed strategy probabilities.*

appendix. Some model specifications show a statistically significant but economically insignificant change (sometimes in the opposite direction as expected).

It’s important to note that the difference in expected value and strategies used between player groups playing against each other are endogenous and the result of some learning process. As such, only the *Increased Price* column can be used to identify our natural experiment. That said, players of all types are highly aware and adjust their strategies to large differences in player skill.

The reason why high skill players are unaffected by the price change cannot be explained by some clever adaptation to the game structure. The lack of change in expected value seems instead to come from differences in play style. High skill players tend to play more aggressively and selectively, lending itself to shorter hands and fewer actions after a flop is dealt. High skill players also win faster against ephemeral players (the largest losers in our zero-sum economy) leading to shorter matches and less rake collected.

Overall, these results indicate that low skill poker players do not respond to changes in payoffs, even large ones, in our game. These results cannot be extrapolated to high skill players because that group was not affected by the price increase.

1. **Rake, expected value and strategy**

We now turn our attention to how much players strategies *should have* adapted.

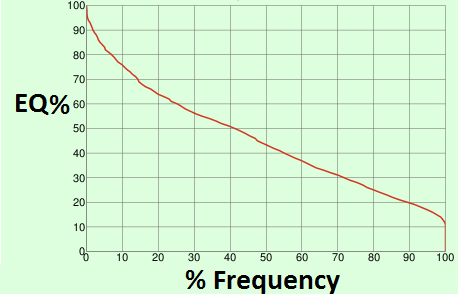
We can decompose the expected value of a poker position as such:

Where is the expected value of the situation (or information set) for player *P1* against player *P2*. *P* is the amount of contested money (denominated by the minimum bet, or big blind, in the game). is the *“equity”* of player 1’s private holding or distribution of possible private holdings against the distribution of private holdings by player 2. The “equity” is the percentage of times player 1 will win against player 2 if no further actions are taken by either player until the hand ends). If a player wins the hand, his payoff is , that is the half of that the opponent contributed. Similarly, if the player loses the hand, his payoff is . The variable is a non-negative real number, and is a free parameter to adjust outcomes for players’ strategies in the situation . It can be thought of as the “equity realization factor”. In the case that both players play in such a way to have the same outcome as not acting, would take the value of 1. Lastly, is the price charged, that is the amount of money removed from *P* by the host of the game. Note that it depends on the situation as the price is not charged before a situation where a flop is dealt.

We are interested in the effect change in on the expected value of a situation, and by consequence his strategy.

Because situations we are interested in have a distribution of possible holdings for both players, is a distribution rather than a scalar. This for the player second to act is discrete, containing all possible private holdings in poker because it is the player’s first decision node.

**Figure 5** plots in the *second to act facing raise* situation when facing a low skill player as defined in section 4. The population of low skill players elect to raise with a 70% frequency, which we model as the strongest 70% of possible private holdings. The player second to act facing raise will have 100% of possible holdings as this is his first action. This is done by 100,000 montecarlo trials simulating a poker hand with no actions taken by either player using the *Odds Oracle* software[[7]](#footnote-7).



**Figure 5.** *Emprirical distribution of for the “second to act facing raise” position when facing a “low skill” opponent in the treatment group. X axis is the frequency at which the player has the in the Y axis.*

Using the theoretical framework and empirical values, we can estimate the best strategy under the new price for the low skill group as defined in section 3 by the k-means algorithm. Because the situation is repeated for all 1326 possible, the total expected value of electing to continue in the information set is:

This is the sum of the expected values in (2) for all possible private holdings of the player at the decision node. is the probability of having the private holding in the information set .

The values of are 0 under regular price and -0.711 under the increased price, as seen in section 6.1. The average observed in the *second to act facing raise* situation is 4.93minimum bets**.** If players do not adjust their strategies, which is empirically the case, is going to be held constant under a change in . Assuming is the same for all , the average for the population at this decision node is 0.86. This is done by estimating the population for the situation by montecarlo simulations (0.585), assuming players continue with the best 28% of private holdings when ranked by against all possible hands and solving for . is estimated to 0.178 before the price increase and 0.290 during the price increase.

Using these values, we can identify all private holdings which have expected value less than -1 as holdings which a player should elect to fold with instead of continuing. This works out to

continuing with a 5.3% lower frequency.

1. **Conclusion**

We exploited a natural experiment (increase in the price to online poker) to study player behavior and test the credibility of models of learning in games.

We find that low skill poker players do not adapt their strategies at all in response to a significant change in payoffs. High skill poker players were effectively unaffected by the change in price due to particularities of the pricing structure and their aggressive approach to the game, leading to a small to nil effect from the price increase.

These results imply that any model of learning which bases its update method on payoff information (eg. reinforcement learning or regret minimization) is unlikely to explain player behavior well in large games with noisy payoffs like the one studied here.

**References**

Babichenko, Y. and A. Rubinstein. 2016. “Communication complexity of approximate Nash equilibria.” *Arxiv preprint*. Arxiv: 1608:06580, 2016

Bronowski, Jay. 1973. “The ascent of man,” Documentary. Episode 13

Bonhomme, S., and E. Manresa (2015): “Grouped Patterns of Heterogeneity in Panel

Data," Econometrica, 83(3), 1147-1184.

Bonhomme, S., T. Lamandon and E. Manresa (2017): “Discretizing unobserved heterogeneity," Working paper.

Camerer, C. F. 2011. “The promise and success of lab-field generalizability in experimental economics: a critical reply to Levitt and List.” SSRN Working Paper

Camerer, Colin, and Teck-Hua Ho. 1999. “Experienced-weighted attraction learning in normal form games." *Econometrica*: 827-874

Hastie. T., R. Tishrani and J. Friedman. 2009. “Elements of Statistical Learning*”.* Springer-Verlarg, 2nd edition.

Johansen, Michael. 2013. “Measuring the size of large No-Limit games.” *Arxiv preprint*. Arxiv: 1302:7008, 2013

Levitt, S. and J. A. List. 2007. “Viewpoint: On the Generalizability of Lab Behaviour

to the Field.” Canadian Journal of Economics 40(2): 347-370

Levitt, S.D., List, J.A. and Reiley, D. H. 2010. “What Happens in the Field Stays in the Field:

Professionals Do Not Play Minimax in Laboratory Experiments.” *Econometrica 78:1413- 1434*

Lin, C. C., and S. Ng (2011): “Estimation of Panel Data Models with Parameter Heterogeneity

when Group Membership is Unknown”, to appear in Journal of Econometric Methods

Linnet, J. et al. 2010. “Experienced poker players differ from inexperienced poker players in estimation bias and decision bias.” *Journal of Gambling Issues* 24

Fudenberg, D. and D. Levine. 2016. “Whither Game Theory? Towards a Theory of Learning in Games,” Journal of Economic Perspectives, 30, 151–170.

Hart, S. and A. Mas-Colell. 2000. “A simple adaptive procedure leading to correlated equilibrium.” *Econometrica*, 68:1127–1150

Hart, S. and A. Mas-Colell. 2003. “Uncoupled dynamics do not lead to Nash equilibrium.” American Economic Review, 93:1830–1836

Mohlin, E., R. Ostling and J. T. Wang. 2015. “Learning by Imitation in Games: Theory, Field, and Laboratory.” Oxford discussion paper series.

M. Moravcík, M. Schmid, N. Burch, V. Lisý, D. Morrill, N Bard, T. Davis, K. Waugh, M. Johanson, and M. Bowling. 2017. “Deepstack: Expert-level artificial intelligence in no-limit poker.”

ArXiv e-prints

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: 164-212.

1. Source: pokerscout.com reports [↑](#footnote-ref-1)
2. Pokerstars was acquired in June 2014 by the Quebec firm Amaya inc. The upper management of Amaya inc., including its then CEO David Baazov and 10 co-conspirators, were later indicted of insider trading by the Quebec financial markets regulator for this transaction (AMF filing 2016-11). If knowledge of the change in management was worthy of insider trading, it is reasonable that the new management and its expected decisions were unforeseen from the poker players’ point of view. [↑](#footnote-ref-2)
3. The structure of the game of poker admits only a mixed strategy Nash Equilibrium. Any skilled poker player will play a mixed strategy in commonly encountered situations, even off equilibrium. Pure strategies leave players too easily exploited by response strategies. [↑](#footnote-ref-3)
4. https://www.pokerstars.com/en/blog/corporate\_blog/2015/2015-rake-rollback-153140.shtml [↑](#footnote-ref-4)
5. See <https://www.pokertracker.com/guides/PT3/general/statistical-reference-guide> for definitions. The statistics are the frequency at which a player willing plays a hand, the frequency at which a player raises when first to act, the ratio of aggressive to passive actions, the ratio of hands played when second to act, the frequency at which the player re-raises when second to act, the frequency at which a player makes a “continuation bet”, the frequency of “donkbets” on flop, the frequency of folding vs. a flop bet, and check raise frequency son flop. [↑](#footnote-ref-5)
6. See (DOYLE). For instance, in the Nash Equlibrium solved in (AU 2014) for the related game of limit hold’em a player takes this action with <0.5% frequency in his aggregated strategy (see http://poker.srv.ualberta.ca/preflop) [↑](#footnote-ref-6)
7. http://www.propokertools.com [↑](#footnote-ref-7)