**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. Controlling for unobserved heterogeneity, we find that*

*players respond more to other players’ behavior than to changes in payoff.*

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 unevenly through the poker game tree, the price change affects the expected value of some strategies and not others. As such, it serves as a rare instance of a natural experiment in behavioral game theory outside of the usual university laboratory environment.

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 make the analysis tractable. No-limit Texas Hold’em, the game played in our dataset, stands in sharp contrast to this trend with an approximate 10140 possible situations (Johansen, 2013).

On one hand, this makes many usual avenues of analysis impossible – simply indexing possible strategies can be intractable. On the other hand, 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, referring to poker, once stated: *“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).

There has been debate over the external validity of lab experiments in economics (see for instance Levitt & List (2007) and the response by Camerer (2011)), generally contrasting the stereotypical low-stakes, artificial experiments on students in academic labs with higher stakes field settings with higher skilled agents. There is agreement that a complement of field and lab experiments contribute more than the sum of their parts, however. Regardless of one’s opinion, natural field experiments like the one found in our dataset are something of a rarity in empirical game theory.

**Learning in games**

Fudenberg & Levine (2016) review the state of game theory and argue that models of learning and social preferences are the important avenues of future research. Specifically, they note that

the predictive power of learning models depends on circumstances, and that we do not have a good understanding of which performs well when.

This problem is made stronger considering theoretical results. Hart & Mas-Collel (2003) show that no uncoupled dynamics – adjustment rules that consider only one’s own payoff function – can lead to Nash equilibrium in all games. Furthermore, Babichenko & Rubinstein (2016) no learning rule using information from previous games will converge efficiently to -Nash equilibrium for every possible game.

The two most popular models of learning in repeated games are belief learning and reinforcement learning. 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. Fitting a EWA model to full scale poker in our dataset is intractable.

One learning rule which has been outstandingly effective when applied to poker is *regret matching*, introduced by Hart & Mas-Collel (2000), where a player plays a strategy proportional to the counterfactual accumulated payoff had he played his other possible strategies in past rounds. 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 (LIBRATUS).

Instead of myopically considering the action taken, the player can reflect ex-post of the counterfactual performance of his other options. Because of this, regret matching can be considered a more realistic update rule than reinforcement learning in games with many possible strategies, most actions will never be reinforced under simple reinforcement learning. For example, Mohlin, Ostling and Wang (2015) use this argument to demonstrate that simple reinforcement learning cannot explain empirical behavior in field data.

Both regret matching and reinforcement learning models can be classed under **payoff-dependent learning rules,** where a player updates his strategy based on payoff information. On the other hand, belief learning, as well as other update rules like imitation learning can be classed as **social learning rules**, where the update method is based on observed behavior of other players. EWA models live in the intersection of both classes of learning rules.

The paper is organized as follows. In **section 2**, we discuss the data and the 2014 price change. In **section 3**, we discuss unobserved heterogeneity in poker players and methods to deal with the problem. In **section 4**, we discuss the game structure of poker and the effect of the price change on said structure. In **section 5**, we present empirical results in the change in value of decisions for players from the price change. In **section 6,** we present empirical results on the change in behavior from players. **Section 7** concludes.

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

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

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 (*without information loss). 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.

1. **Unobserved heterogeneity in poker players**

Like other popular table games, poker has a population of highly skilled professional players. In game playing artificial intelligence, Moravcik et al. (2017) and Brown and Sandholm (2017) decisively test the performance of their poker agents against professional human 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.

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 and are effectively 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, for example in Lin & Ng (2012), Bonhomme & Manresa (2015) and Bonhomme, Lamandon & Marsenna (2017). These 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.

We use the clustering algorithm on 15 variables of interest: 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 pulled from *PokerTracker[[5]](#footnote-5)* of strategic behavior at the table: VPIP, Attempt to Steal, AF, Fold BB vs. SB raise, 3bet, Cbet flop, donk flop, fold vs. flop cbet, and check raise flop.

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.

The last method is to statistically partition the dataset between players who are likely professional players by a t-statistic. 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 2** plots the hands played and winnings in US dollars for the treatment group in our dataset. Purple players are professional as defined by the statistical test above. In **figure 3,** 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.

**Figure 3** demonstrates the change in players’ hourly wage by group. We can see that high skill players are unaffected by the change in price, while low skill players in the treatment group and ephemeral players are heavily affected. We can also see that

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.** 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 “facing raise”, by player 2 facing player 1’s initial raise. It is the node where the choice between strategies has the largest change in the expected value from the price change 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 taking this position in a panel model with fixed effects:

(1)

Where 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 are found in *X.* Controls based on time (time of day, day of the week, etc.) were found insignificant in all specifications.

**Table 1** plots 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 notice that there is a large effect on low-skill regular players, and no significant effect on high skill players. Ephemeral players are unaffected by the change, even when aggregated as a group.

1. **Strategic behavior**

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

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.

All players call or re-raise significantly less 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 opponent from the treatment group.

The change in price seems to have a small effect on players’ strategies, which is statistically significant, but economically insignificant. Players who are classed as professionals by the t-test described in section 3 react in the opposite direction as expected, for example.

These results imply that players are highly aware and elastic to large differences in player skill, and generally more elastic to differences in opponent behavior than to differences in payoff of strategies.

**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.

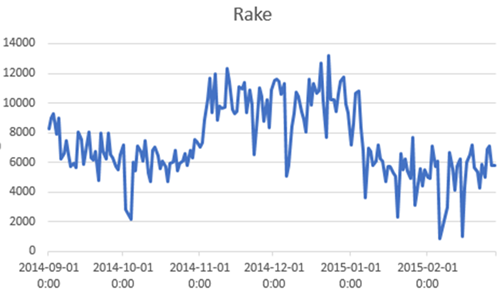
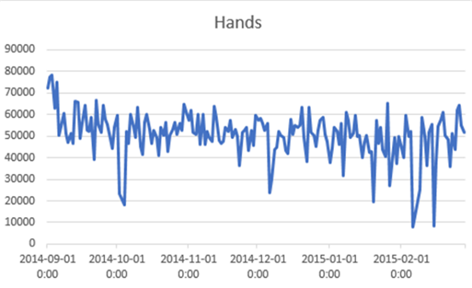
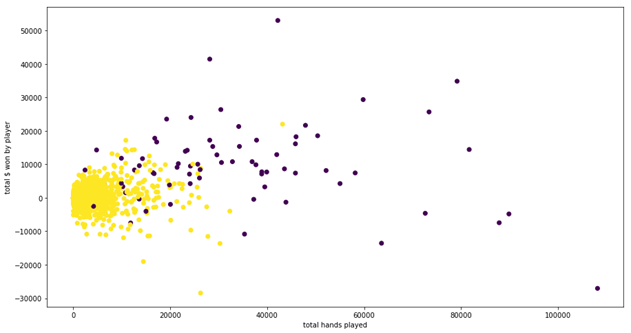
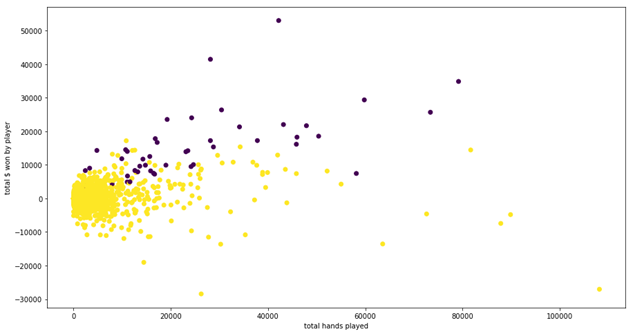
Figure 1

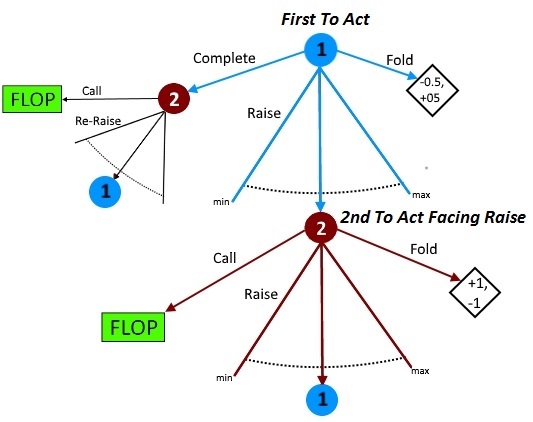
Figure 2.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| BB EV |  | **Constant** | **Price Change** | **High Skill Opponent** | **Ephemeral Opponent** | **Ephemeral**  **Opponent**  **x Treat** | **Low skill**  **x Treat** | **High Skill x treat** |
| **Treatment**  **Group** | **All** | 0.555  (0.124) | -0.647  (0.215) | -1.410  (0.285) | -0.017  (0.190) |  |  |  |
| **High skill**  **(k-means)** | 2.291  (0.321) | -0.340  (0.308) | - 1.567  (0.486) | -1.007  (0.363) |  |  |  |
| **Low skill**  **(k-means)** | 0.059  (0.171) | -0.762  (0.291) | -1.861  (0.362) | 0.020  (0.230) |  |  |  |
| **High skill**  **(t-stat)** | 2.531  (0.426) | 0.212  (0.422) | -0.688  (0.669) | -1.010  (0.478) |  |  |  |
| **Low skill**  **(t-stat)** | 0.287  (0.162) | -0.740  (0.248) | -1.675  (0.316) | 0.003  (0.210) |  |  |  |
| **Ephemeral**  **Players** |  | -3.492  (0.318) | 0.000  (0.000) | 0.369  (0.564) | 0.642  (0.391) |  |  |  |

Table 1

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Fold BB % |  | **Constant** | **Price Change** | **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



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. 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 [↑](#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)