Esport: Suspense and surprise, luck and skill

By John Norell, april 2018

I'm in the middle of writing my bachelor thesis in economics right now and so far it's been a roller coaster and we've had to change subject multiple times now, which is a bit stressful to say the least. Since I've spent a lot of time reading research and thinking about what could be potentially interesting research questions, I figured I'd share a little bit of what I've learnt and perhaps someone in the future will benefit from it.

Suspense and surprise in video games

The first idea I had of what would be an interesting subject to delve into was suspense and surprise in e-sport. I came across a theoretical article published in 2015 where the authors develop a microeconomic model for quantifying suspense and surprise[1]. (I first heard about it from an episode of the podcast freakonomics[2]).

The starting point for the suspense and surprise model is thinking about why it is that people watch sports, read mystery novels, follow political elections and a range of other activities that wouldn't really make sense if it was the acquiring information that was the purpose. A more reasonable explanation would be that people get joy from the way the information is delivered in these situations. Basically, their theory is that people gain utility from being held in suspense or the feeling of being surprised. Researcher have in experiments shown that suspense do increase the entertainment value of sports[3].

To try and explain it simply, suspense is what you experience when you anticipate that soon something big will happen that will change your beliefs. A situation with high suspense could for example be that you think that the outcome of the serve is likely to determine the outcome of a tennis match, or that the who gets the votes from a particular wing state is likely to determine the winner of a political election. In the model, they express this

---

[1] Ely, J., Frankel, A., & Kamenica, E. (2015). Suspense and surprise. *Journal of Political Economy*, *123*(1), 215-260.

[2] The episode was published 29 july, 2015. It can be accessed on
http://freakonomics.com/podcast/how-to-create-suspense-a-new-freakonomics-radio-episode/

[3] For example Bryant, J., Rockwell, S. C., & Owens, J. W. (1994). Buzzer beaters" and "barn burners": The effects on enjoyment of watching the game go "down to the wire. *Journal of Sport and Social Issues*, *18*(4), 326-339.

mathematically: suspense is then measured by the variance of your beliefs in the next time period.

Surprise is then measuring how much your beliefs changed from the previous time period to this one. If something happens that totally changed your beliefs about will win an election, who's the murderer in the mystery novel or win the football match, then you will be feeling a lot of surprise. These two concepts are related to each other, suspense is about what you expect to happen, surprise is about what then actually happens. Both suspense and surprise is about how information is delivered to you, not the actual information.

The problem with this theory is that it's a bit tricky to test. This is because we need data on people's subjective beliefs about the likelihood of specific events and how these change with time when more information is revealed. One way to deal with this is to use live betting odds (which measure the aggregated beliefs of the bettors about the likelihood of an event happening) as an approximation for people's subjective beliefs[4]. This approach was used by (Bizzozero et al, 2016)[5] to apply the suspense and surprise model to tennis matches at Wimbledon.

Betting odds is enough to create estimations on the level of surprise the audience experience at any given point during a sport match (or any event as long as there are live betting markets operating). If you have data on concurrent viewership (how many people watch the game at any given point of the match), it's not very difficult to test if there is a relation between surprise and viewership. the idea being that people are less likely to stop viewing the match (turn off the tv) when something surprising is happening or has happened.

Suspense is a bit more difficult to quantify, since you not only need people's actual beliefs about the likelihood of player X winning the match at every point of the game (for this we can use the live betting odds again), we also need to know people's beliefs in hypothetical cases. What would their beliefs have been if player Y would have won this serve instead of what actually happened, that player X won it. And we need this for every scenario at every point of the game. This is what is meant by the variance of beliefs in the next time period. If there are

[4] See for example Wolfers, J., & Zitzewitz, E. (2004). Prediction markets. *Journal of economic perspectives*, *18*(2), 107-126.
[5] Bizzozero, P., Flepp, R., & Franck, E. (2016). The importance of suspense and surprise in entertainment demand: Evidence from Wimbledon. *Journal of Economic Behavior & Organization*, *130*, 47-63.

only two outcomes of a tennis serve, then suspense is measure by the distance between beliefs based on what happened and what you would have believed in the counterfactual case.

The research article I mentioned above where the authors look at suspense and surprise at Wimbledon, they create a simulation model to create predictions about the likelihood of each player winning based on data of what happened in the game as well as historical data about each players previous performances. The predictions by the simulation model was very similar to what the live betting odds predicted, and by tweaking this model they were able to make predictions about the counterfactual situation in every point of the game and hence estimate the level of suspense at every point in the game. To be honest, that is a very clever way to do it, I was impressed reading their method section. For more complex games, where the game can't be easily divided into subgames and turns but instead play out continuously, it's much more difficult to create simulation models and estimate suspense. In video games such as dota, CS:GO or Fortnite, the number of possible actions at every point in the game is almost endless so quantifying suspense in such a game would be quite a daunting task. Though I actually found a research plan from a group of researchers who planned to do this to see if the level of suspense in a match in the video game League of Legends can be used to predict the likelihood of players continuing to play the game or not.[6] Unfortunately I could not find if they actually went through with the plan or not.

As is evident from the statistics on viewership on video game streaming sites such as Twitch [7] and youtube gaming, entertainment habits are changing rapidly, especially among young people. The audience of people playing video games is growing fast and this creates interesting questions about what it is that makes a video game popular with the audience. Video game developers are no longer only in the industry of providing entertainment for the players, but also entertainment for a growing streaming audience. If the streaming viewers and the gamers have different preferences then that pose an interesting question for game developers how to weigh there interests. Even if the game developer don't make much money on the viewers directly, it's still a form of marketing for them.

If the audience have a preference for suspense, then we can expect the strategies used by popular streamers to be different from the average players because they have a crowd to

[6] Lance, B., Zhang, W., Santiago-Calderon, J., Bjurstrom, S., & Crisanti, C. League of Legends: Testing the Effects of Suspense, Surprise, and Friendship on Satisfaction and Entertainment.
[7] The growth of Twitch since the start in 2011 can be seen on https://twitchtracker.com/statistics

entertain. If we can identify how these strategies differ, then that can tell us something about the audience.

Two of the most popular games right now on twitch are *PUBG* (short for Player Unknown Battlegrounds) and *Fortnite*. These games are both what are called *battle royal games* which are similar to the tournament in the hunger games series for the one who've read the popular books or seen the movies. A hundred people are dropped on an island to scavenge for weapons, armour and other useful equipment. A player wins the game by being the only survivor left on the island. It's possible to win this game by staying hidden and wait for all the other players to kill of each other. Such a player strategy isn't very suspenseful to watch for the audience of a video game streamer however. A players who throughout the game continuously risk being killed by engaging with enemy players will produce a more suspenseful game for the audience.

I suspect therefore streamers are more likely to apply more aggressive and/or riskier strategies compared to non-streaming player. Can this be measured? I think so, if more aggressive tactics also mean they will encounter more enemies, then they are both more likely to be killed and to kill more players than other players that are as skilled as them. I have not however found that kind of statistic, but maybe someone will be able to find or gather that in the future.

Skill and Luck in E-sport

Due to the complexity of the method and our problems of finding data we changed course, to instead look at the role of luck in E-sport tournaments. As I've described above, surprise and suspense are concepts related to the way information is revealed. In order for there to be suspense in a match, there need to be a level of uncertainty about the outcome. So then the question becomes, where does this uncertainty stem from? When you are watching a football match, this uncertainty I believe can come from two sources. First from you not knowing which is the better team. If a game of fotball always resulted in the victory of the more skillful team, then any suspense or surprise must come from uncertainty about which is the better team. The other possible source of uncertainty comes from the fact that the game doesn't perfectly measure the relative skill level of the players. Sometimes the better team loses a game, so there is a component of randomness in outcomes. This component we can call "luck". Based on this idea, a game/sport/activity where luck have no part whatsoever,

and where the skill of participants/players/teams doesn't change much from game to game, would be very predictable. And even though there might be other reasons for observing this game, it wouldn't be very suspenseful.

In the book *"the success equation"*, Michael mauboussin[8] explain that while success in some activities are determined solely by chance (such as the lottery) and some solely by skill (playing a musical instrument for example), for most things in life both skill and luck have part in the outcome.

There is a lot of literature on the role of skill vs luck in different activities, mostly these researchers have studied games such as poker and chess. These papers often relate to the discussion of weather or not a game shall be classified and therefore also regulated as games of skill or of games of luck. Skill here is understood as to what extent the individual or player can influence the outcome of the activity or the game[9]. One conclusion to draw from these studies is that while roulette and chess are easily classified, most games are not.

Instead of asking whether of not a game is a game of skill or of luck, a better question if *when* the game becomes a game of skill. If a player can gain even the slightest advantage by skill, then it will become a game of pure skill when playing an infinite number of games. So even the game of Rock-Paper-Scissors becomes a game of skill if you play enough matches[10].

In games where players compete against other players, skill is no longer an absolute characteristic but a relative one. What I mean by that is that how much of an advantage you get from your abilities depend on the abilities of your competitor. In professional sport leagues where all players are very skilled, the variance of skill is low. This causes luck to be more important to succeed. This is quite counter intuitive, that as everyone gets more skilled in the game, luck will have a greater influence on the outcome. So while many game becomes more of a game of skill when you play more matches, it becomes more of a game of luck when everyone gets more skilled.

---

[8] Mauboussin, M. J. (2012). *The success equation: Untangling skill and luck in business, sports, and investing*. Harvard Business Press.
[9] Dreef, M., Borm, P., & Van der Genugten, B. (2004). Measuring skill in games: Several approaches discussed. *Mathematical methods of operations Research*, *59*(3), 375-391.
[10] Fiedler, I. C., & Rock, J. P. (2009). Quantifying skill in games—Theory and empirical evidence for poker. *Gaming Law Review and Economics*, *13*(1), 50-57.

Since we know how randomness behaves, we can measure to what degree skill and luck respectively determine success for different activities. In the book mauboussin does this for the five biggest sport leagues in America (baseball, basketball, soccer, ice hockey and american football)[11].

Can this method be applied to E-sport tournaments? How much influence does randomness have on the outcome of these tournaments basically. The video game industry change so rapidly (some of the most popular games that are streamed at twitch, PUBG and Fortnite, hadn't even been released a year ago) and both the audience and the players seem to be less loyal to one specific game or brand if you compare to traditional sports. For this reason, perhaps e-sport tournaments cluster together around the optimal level of luck-level on this luck-skill-continuum. this optimal level would then constitute the optimal level according to the audience. That is of course only a thought, perhaps e-sports are just as scattered as other sports. If this optimal level of luck involved in the outcome is related to the audience preference for suspense and surprise, or some completely different mechanism, I don't know.

Just to test out the idea I tried it out on the ESL *CS:GO 5on5 Open League Spring 2017 Europe* [12]. 422 team of five players joined the group tournament and each team played up to 7 matches, from which top advanced to an elimination tournament. Below you can see a histogram of all teams that played 5 matches and how many wins each team won. If the relative skill level of the competing team of each match would perfectly determine who won the match, then we would expect a flat distribution. While if each match was determined solely by randomness, say a coin flip, then we would expect a binomial distribution. what we see below is something in between, which suggest there is both skill and luck involved here.
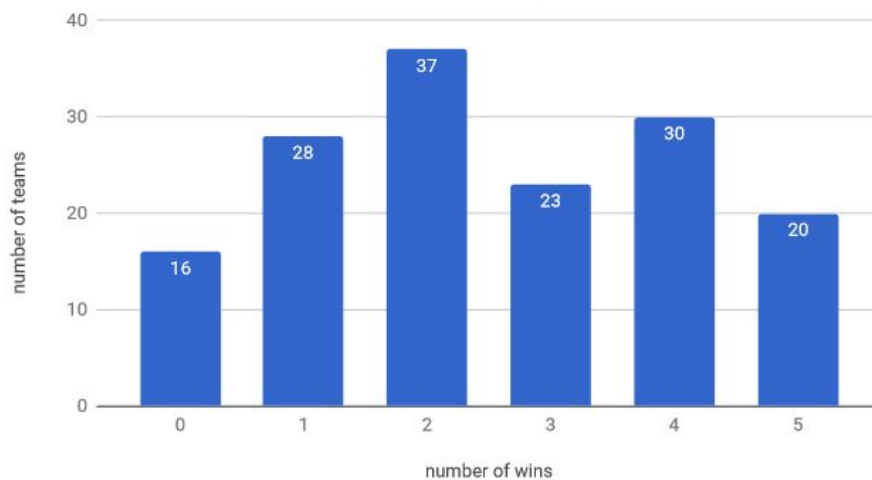
---

[11] A summary of these findings can be seen in this short video, *"Why underdogs do better in hockey than basketball"*, published 5 june 2017 by Vox: https://www.youtube.com/watch?v=HNlgISa9Giw
[12] *The statistics is found at*
https://play.eslgaming.com/counterstrike/csgo/csgo/open/league/spring-2017/rankings/
There are data for 9 tournaments of this type, going back to 2014 according to their website

Number of wins for all teams who competed with five matches

Next I calculated the variance of luck and subtracted that from the total observed variance to estimate the variance of skill[13]. When the number of matches a team play increase, the role of skill increase, but it's around 50% for teams playing 5-6 matches each[14]. The problem though is that I've not been able to find any good peer-reviewed sources for this method[15]. For this reason, we had to change approach again, this time looking into happiness research, by suggested from our supervisor. I can't say I'm super excited, but it's only a bachelor thesis, so I'll manage.

---

[13] The method to estimate the level of skill and randomness in the outcome of sports is explained here: http://blog.philbirnbaum.com/2006/08/finding-true-talent-level-for-outcome.html
Applied to measure the level of skill and luck in the NFL sports league is found in three blog posts here:
http://archive.advancedfootballanalytics.com/2007/08/luck-and-nfl-outcomes.html
http://archive.advancedfootballanalytics.com/2007/08/luck-and-nfl-outcomes-2.html
http://archive.advancedfootballanalytics.com/2007/08/luck-and-nfl-outcomes-3.html

[14] A problem here is that all teams haven't played the same number of matches, because some teams leave the tournament when they lose a couple of matches, and some don't show up at all. So we divided them based on matches played. Then we took out the variance for each group and calculated the variance of "luck"/chance. What we found is that the more matches teams play, the more skill is likely to impact the outcome. var(skill)/var(observed) for matches when teams play 7 matches is 0,69, 0,50 for 6 matches, 0,48 fr 5 matches, 0,45 for 4 matches and 0,36 for 3 matches.

[15] A related approach is used in this article that looks at if people betting on soccer accurately take the reversion to the mean into account in their predictions. Lee, M., & Smith, G. (2002). Regression to the mean and football wagers. *Journal of Behavioral Decision Making*, *15*(4), 329-342.