Strategy and Incentives in Multi-stage Tournaments

Christopher Liu

Abstract

Tournaments divided into different stages, such as a round-robin stage proceeding to an elimination stage, present opportunities for anti-competitive strategic behavior. Specifically, in certain situations teams may be incentivized to lose, indicating that such tournaments are not monotone. We developed a Monte Carlo simulation of multi-stage tournaments in order to study the prevalence of such situations, as well as the impact they may have on the overall tournaments. We found that strategic behavior may be incentivized in a moderate proportion of situations. Moreover, when such incentives exist, we confirmed that teams that decide to intentionally lose do indeed improve their chances of winning the overall tournament. Lastly, the presence of such incentives have minor effects on the competitiveness of the overall tournament.

1 Introduction

In the 2012 London Olympics, the world was shocked by the disqualification of four women's doubles teams in Badminton [1]. All of the teams involved, which included the top-ranked duo in the world, had already qualified for the quarterfinals; apparently, they were trying to avoid matching against the already-qualified number 2 pair in quarterfinals, another Chinese team. As such, they each attempted to intentionally lose their respective preliminary matches, amidst booing from stadium crowds. While most commentators denounced the teams for their anti-competitive actions, some noted that their actions were incentivized by the tournament structure. In this paper, we aim to empirically study this phenomenon.

1.1 Problem Definition.

In order to perform our study, we specify a specific tournament structure, taken from the 2021 League of Legends World Championships. The tournament involves 16 teams, and a group stage proceeding to a knockout stage. The teams are divided into 4 groups of 4 teams, with each group playing a double round robin. The top two groups from each stage are seeded into the knockout stage, which is a single elimination tournament.

Round Robin. A round robin tournament is one in which each competitor meets every other competitor. In our case, they meet each competitor twice. There are various ways of scheduling such matches, we simply use an arbitrary schedule and repeat it twice.

Single Elimination. A single elimination tournament is one in which the losing competitor in each match-up is eliminated from the tournament, with the winners of each round proceeding to the next. If the total number of teams is a power of 2, as in our case, then no byes are required and each remaining team plays in each round. Whereas group stage matches are best-of-1, knockout stage matches are best-of-5, requiring 3 game wins to win the overall match.

Seeding. Both the group stage and knockout stage are affected by seeding, which determines the position of each team in their brackets. Seeding therefore affects the teams they match up against and the order in which they do so. Seeding is typically done with the goal that stronger teams are less likely to meet until later in the tournament to improve overall competitiveness. We only focus on seeding in the knockout stage for this paper.

Seeding in the knockout stage works as follows: first-place teams from each group, which we call the upper bracket, are matched up against second-place teams, which we call the lower bracket. Moreover, teams from the same group cannot be on the same side of the overall bracket, meaning that they cannot meet until the finals.

The focus of this paper is on this first rule for seeding. If teams believe that second-place teams will be stronger than first-place teams, then they may be incentivized to themselves attempt to place second in their group rather than first. In essence, this represents a violation of *monotonicity*, that winning a match should only ever improve your expected outcome. This is what occurred in the 2012 Olympics, as all of the teams involved wished to avoid matching against the 2nd-rated Chinese team, who earlier ended up being placed into the bottom half of the draw.

To study this phenomenon, we perform a Monte Carlo simulation of the specified tournament structure, noting cases in which the lower bracket appears to be stronger than the upper bracket by the time the last group plays, thus shifting the incentives of the group to aim for second place rather than first. If this would alter the overall outcome of the group, then we simulate both variations in the knockout round.

The results of our simulation show that these altered incentives are fairly common in this tournament structure. Moreover, teams that do intentionally throw games when incentivized are improving their odds of overall victory in the tournament compared to teams that always try to win; thus, this tournament structure is not monotone. Finally, in cases where teams are incentivized to throw and doing so would actually alter the outcome of the bracket, we find that doing so does not appear to reduce the "competitiveness" of the tournament, which we define as the average "skill level" of the tournament winner, and indeed it may even have the opposite effect.

2 Background and Related Work

There has been ample amounts of research done on tournament design and incentives. The empirical study of Krumer et al. [2] investigates the effects of scheduling on round robin tournaments, finding that contestants competing in the first and third matches have an advantage over those in the second and fourth in that group. Our study examines another aspect of scheduling, as only later-playing groups can have enough information to be incentivized to aim for 2nd place, as it requires having knowledge of the outcomes of other groups.

In another theoretical study done by Pauly [3], a theoretical framework is developed to analyze the kinds of round robin subtournaments discussed in this paper. He finds that monotonicity captures the notion of strategy-proofness of such tournaments. Pauly finds that for round robin subtournaments with at least 2 groups of at least 3 teams, no tournament design can be symmetric, non-imposed, anonymous, independent of irrelevant alternatives, and monotone. Put more simply, any monotone and thus strategy-proof design must be asymmetric and thus unfair. Our study focuses on the impacts of non-monotonicity on a standard tournament format.

3 Methodology

In order to study this topic empirically, we developed a Monte Carlo simulation of the specified tournament structure, noting cases in which teams are incentivized to intentionally lose games, as well as simulating instances where such teams do and do not intentionally throw their games.

3.1 Data

For our simulation, we randomly generate teams and randomly place them into group brackets. For simplicity, we condense the characteristics of teams into a single value we dub their ”win rate”, which generally represents their skill level. This value is drawn from a normal distribution in the interval $[0, 1]$, with a mean of 0.5 and a standard deviation $\sigma = \frac{1}{6}$.

3.2 Implementation

Team. As mentioned above, each team maintains a ”win rate” value. While we leave open the possibility of implementing individual match-up tables for each team, for the purposes of this study we use a uniform win rate. For each match, the chance of either team is the average of their win rate and their opponent's win rate subtracted from 1.

Round Robin. Each round robin group includes 4 random teams. For the purposes of this study, we did not attempt to seed the groups. Each group plays its matches in a set order in each round, in which each team plays once against each other team in the group. The 4 groups play sequentially, with each group playing out one round before running through the cycle again.

To identify instances in which teams are incentivized to throw games, we focus on the 4th-playing group. Other groups would need to perform some reasoning regarding the outcomes of future groups in order to determine their incentives, which we felt was impractical to simulate and less likely to occur in an actual competitive environment. Thus, during the 4th group's 2nd round, we average the win rates for teams placed into the upper and lower knockout brackets. If the latter's average win rate is higher, then we can conclude that teams in the 4th group are incentivized to aim for 2nd place rather than 1st.

We then compute whether, given this incentive, any team in the group will be incentivized to throw their game. In order for this incentive to exist, they must already be within the top two teams of their group, and must be certain that losing their match will not knock them out of the top two; in other words, they must be certain that even after throwing their match they will still make it to the knockout stage. We first simulate the first 4 matches in the group as usual. Then, for the second to last match, we check if either team is both in the top two and that, if they lose their match, they will not be knocked out of contention depending on the outcome of the next match. For the last match, we check if either team is already in the top two and whether, if they lose their match, the other team may kick them out of contention. If any of the teams match these conditions within the last two matches, then we mark that instance as one in which there exists incentive to lose. This process is described in Algorithm 1. One caveat to note is that we do not explicitly handle tiebreakers in our simulation and simply decide them arbitrarily; as a result, when determining whether teams are incentivized to throw, we ensure that they are avoiding ties as well.

While we simulate the group as if teams play normally, if there is incentive to throw, we also compute whether doing so will alter the outcome of the group. This only occurs if a team that would otherwise place first in the group decides to throw and instead places second, as second place teams would only be securing their placement and thus would not produce a different outcome. In cases where such alternative outcomes exist, when simulating the elimination round we additionally simulate this alternative bracket.

Algorithm 1: Determine if an instances has incentive to throw

```
Result: Returns true if any team the last group is incentivized to throw their last match
Play out the first 7 rounds as normal, as well as the first 4 matches of the last round;
if Average lower bracket win rates are higher than upper bracket then
    if Either team in the 2nd to last match is currently qualifying and, if they lose, cannot
      be knocked out of contention due to results of the next match then
        return true;
    else
        if Either team in the last match is currently qualifying and cannot if they lose be
          knocked out of contention by their opponent then
            return true;
        end
    end
else
    return false;
end
```

Elimination. In constructing our elimination round brackets, we seed upper bracket teams into the bracket, then randomly assign lower bracket teams to match up against them. We do not restrict teams from the same group to being on opposite sides of the bracket as is the case with the League of Legends World Championships format, as that is unlikely to matter for our purposes, but we do prevent teams from being matched up against other teams from the same group in the quarterfinals. We take the winners of each round and seed them in order to the next round, until we arrive at an overall winner.

4 Results and Analysis

A consequence of the nature of our study is that the statistics we keep track of are somewhat difficult to clearly classify. We refer to instances in which the teams of the 4th group are incentivized to place 2nd and in which at least one qualifying team can afford to throw their match and still secure entry to the knockout stage as "instances with incentive to throw". We refer to instances in which the incentivized team throwing their final match changes the outcome of the group as "instances with alternative brackets", with the bracket in which said team does lose their last match as the "alternative bracket". Meanwhile, the original bracket in which that team plays normally and achieves a different outcome (i.e., wins their match) is known as the "non-alternative" bracket, an admittedly confusing bit of nomenclature.

Total iterations	10,000,000
Instances with incentive to throw	2,453,799
Instances with alternative brackets	1,039,909
Average win rate of winners	0.6692
Average WR of winners in instances with incentive to throw	0.6641
Average WR of winners in alternative brackets	0.6626
Average WR of winners in non-alternative brackets	0.6620
Total wins of teams incentivized to throw	349,561
Proportion of instances with incentive to throw	0.1425
Total wins of throwing teams in alternative brackets	148,688
Proportion of instances with alternative brackets	0.1430
Total wins of would-be throwing teams in non-alternative brackets	140,414
Proportion of instances with alternative brackets	0.1350

Figure 1: Results of simulation over 10 million iterations

Caveats. The results of our simulation present a number of notable figures. Firstly, the number of instances with incentive to throw may appear to be overly high; after all, it does not strike as likely that nearly a fourth of such tournaments present such anti-competitive incentives. However, there are a couple of caveats that explain this figure: firstly, this number only tracks instances where teams are incentivized to throw games in order to place second, and thus we are also counting situations where they can safely throw their last game but it will not alter the outcome. I.e., if a team is significantly ahead in first place and thus cannot feasibly place second even if they throw, they are still counted in this figure. Thus, this number may overestimate the number of teams that would actually throw a match.

Even the more conservative figure of instances with alternative brackets, instances in which a team is incentivized to throw, does so, and thereby changes the outcome of their bracket, may strike some as being high. After all, the 2012 Olympics scandal was noteworthy because such events seem to be relatively rare. However, it should be kept in mind that teams in our simulation have perfect information on the win rates of every other team, and are thus able to accurately discern incentives on even very minor margins. Put another way, competitors in the real world are unlikely to throw their match if the margin between the truthful skill levels of the upper and lower brackets are very small, as it is unlikely that they will have sufficient confidence in their judgement regarding optimal play unless the disparity between the brackets are significant. This was the case in the 2012 Olympics, where the 2nd-ranked team in the world placed into the lower bracket. Thus, due to imperfect knowledge of the competitors, they are unlikely to be able to accurately identify incentives to throw unless they are fairly obvious.

Another caveat is that, while in our simulation we determine incentives based on the average win rate of qualifying teams between the upper and lower brackets, it is possible that teams will more heavily weigh the placement of particularly strong teams rather than assigning each team equal weight. In other words, teams may prioritize avoiding the strongest teams over minimizing the average skill level of teams they play against. This is especially true for stronger teams, who may not mind playing against teams of average skill who they expect to win against anyway in favor of avoiding stronger competitors.

It goes without saying that even if teams believe they would be better off losing their match, they may still try to win simply on principle. All of these factors contribute to our figures seeming higher than one would expect.

Competitiveness. We define the competitiveness of a tournament to be its ability to select the strongest team as the winner. Thus, a more competitive tournament will have a higher average win rate for their winners. We can see that the average win rates for winners across various categories of instances are fairly close together. However, we do have to be wary in interpreting these results, as the parameters for selection of sampling has an effect on output. This is evident in the average win rate of instances with incentive to throw being some 0.5% lower than the norm, suggesting that such instances either tend to select less competitive winners, or that such instances tend to have a lower win rate best team. Thus, we cannot directly conclude from this figure that such incentives reduce competitiveness. Indeed, the latter case does make intuitive sense, as instances with a wider range between stronger and weaker teams are less likely to end up with imbalanced brackets, as the stronger teams will be more likely to win their groups compared to in instances where the skill level of teams are closer together.

For the figures of the average win rate of winners in alternative and non-alternative brackets, as the sampled instances are identical, we are insulated from such factors. As such, we are able to use them to draw some conclusions. Notably, the figures are very close together compared to previous figures, suggesting that sampling factors have a larger bearing on average winning team win rate than the exact arrangement of the bracket depending on the strategizing team's actions. However, that the average winning team win rate is higher when said teams decide to throw their games compared to the opposite is interesting. Though on the surface this may be surprising, it does make intuitive sense. Teams are seeded into upper and lower brackets for the sake of improving competitiveness, so as to prevent strong teams from knocking each other out early on in the elimination round. Both figures are sampled exclusively from instances in which, by 8th group stage round, the lower bracket is stronger than the upper bracket on average. The strategizing team is likely the strongest team in its group, as it has the most wins playing normally in the non-alternative brackets. Thus, by throwing and being placed into the lower bracket, we are in essence reversing the upper and lower brackets. As such, from this perspective throwing games can actually *improve* the competitiveness of the tournament.

Outcomes. The last section of our results tallies the total instances in which the team which is incentivized to throw wins the entire tournament. We can see that all three figures present a higher proportion of wins compared to an even proportion of $\frac{1}{8} = 0.125$ for each quarterfinal team.

We can see that, in instances where throwing would make a difference, the affected team improves their odds of winning by throwing, increasing their proportion of victories by roughly 0.8%, a nearly 6% improvement. We have therefore empirically verified that throwing games can indeed be incentivized in the group stage, and that in such cases teams can improve their odds of winning by throwing a match. Thus, the tournament structure discussed in this paper is not monotone.

5 Conclusion

Our simulation supports the conclusion that multi-stage tournaments like the one discussed in this paper are not monotone. In certain circumstances, teams may be incentivized to intentionally lose their group stage matches, and by doing so may improve their overall chances of winning. Thus, viewed from a broader perspective on competition, teams such as those disqualified in the 2012 Olympics are indeed doing their best to win, but in reference to the overall tournament rather than individual matches. It is therefore not a straightforward matter whether we should blame such teams for their actions, as it is explicitly against their interests to win their matches; rather, is it not the fault of the tournament design?

This paper does not attempt to propose or analyze alternative structures; indeed, Pauly's paper [3] suggests that a perfect tournament design may not be possible. Of interest, and something that Pauly also considers, is whether repechage systems such as double-elimination tournaments, in which teams are only knocked out upon losing two matches rather than one, may help mitigate incentives to throw in the group stages. By reducing the importance of each individual matchup, the incentive to attempt to place into upper or lower brackets may become less clear. From a practical perspective, it is also likely that increasing the complexity of the knockout bracket will make teams less likely to identify incentives even if they exist, thereby reducing the likelihood of teams opting to throw their matches.

Our simulation makes a number of design decisions for purposes of simplicity. Some of these have been previously mentioned. We use uniform win rates for each team rather than having specific matchups against different teams. We don't consider rare situations where teams may opt to throw games apart from the circumstances studied in this paper. Lastly, we draw team winrates from a normal distribution, but different leagues may have alternative skill distributions. Informally, running our simulation using different distributions seems to alter the specific values but not their relations.

Tournament design is a field rife for further exploration, and we are excited for what future studies may reveal.

6 Acknowledgement

We would like to thank Shikha Singh for her help and guidance in this study.

References

[1] Ken Belson. Olympic ideal takes beating in badminton. *The New York Times*, 2012.

[2] Alex Krumer and Michael Lechner. First in first win: Evidence on schedule effects in round-robin tournaments in mega-events. *European Economic Review*, 100:412–427, 2017.

[3] Marc Pauly. Can strategizing in round-robin subtournaments be avoided? *Social Choice and Welfare*, 43(1):29–46, 2014.