Comparing the Efficacy and Fairness of the 2024/2025 UEFA Champions League Format with Other Tournament Formats

Gabriel Faes

Abstract

The UEFA has introduced a new format for the 2024/2025 Champions League. We explore how efficacious and fair it is against other tournament formats (with similar characteristics, i.e., similar number of rounds, matches and players) using Monte-Carlo simulations equipped with the Bradley-Terry model. We find that the new format is less efficacious than other similar tournament formats, namely a round robin group stage, but only marginally. Additionally, lower seeded teams have a slightly higher chance of winning which encourages teams to game the system and intentionally lower their country's or club's ranking to obtain a lower seed.

I. Introduction

Globalization has allowed tournaments to tremendously grow in size. Spectators can now watch their favourite teams play popular games no matter their geographical location. With this (still growing) attention, some tournaments now have prize pools which can reach millions of dollars, with some nearing a billion. With so much at stake, the formats of tournaments have started being much more closely investigated since the last century. It is essential for all invested parties (whether organisers, players, online or in-person spectators, or TV broadcasters and more) that the tournament format meets certain criteria, such as being fair, or generally allowing the stronger teams to win. Determining these criteria has been the topic of on-going research; we will study here efficacy and fairness.

A variety of standard tournament formats exist, each created to fulfil different needs. The single elimination tournament is the fastest at determining a winner, while a round robin tournament is much longer but crowns a stronger player more often. Tournaments also have different seeding methods, which are used to place players in the tournament in a certain order which may affect the probability that players win the tournament.

The UEFA has designed a new tournament format for its Champions League for the 2024/2025 season. We focus here on its league phase and knockout/single elimination phase. In addition to changing how the first stage/league phase functions, it has also increased the total number of teams to 38. The bracket/knockout stage remains a standard single elimination bracket (albeit with an additional playoff round for teams to qualify into the round of 16).

To determine the probability that a player wins a given tournament format with a specific seeding position, Monte-Carlo simulations will be performed where the probability that a player wins a match is given by the Bradley-Terry model. These simulations will be used to measure the different tournament formats and seeding methods against the tournament design criteria and measure which ones perform better for different criteria.

II. Terminology

II.1 General Terminology

II.1.a Tournament

A competition pitting players or teams of players against each other using a specific game with the goal of crowning a player or team as the winner of the tournament. As most games/sports are one versus one (and those will be the only games considered here), players and teams will be used interchangeably to refer to a single entity attempting to win the tournament as a whole.

II.1.b Prize Pool

The total prize money available for a tournament. It is allocated in accordance with the tournament rules, with higher placing teams generally winning a bigger percentage of the prize pool.

II.1.c Game, Set, and Match

A match is comprised of a (or multiple) set(s) of games. The player who wins the most games wins the set and, if there is only one set, the match as well; if there are multiple sets, the player who wins the most sets wins the match. The game (game or sport the tournament is about) and a game in a match/set are not to be confused. In football, a match is generally comprised of a single or two 90 minutes games. In the case that a match is comprised of a single game, game and match are used interchangeably.

II.1.d Tiebreaker (in a match)

In the case where a set or match has an even number of games or sets and both players win the same number, a tiebreaker may be called (depending on the tournament rules which may or may not allow a tie in a match). In the case that one is called, it is used to break the tie and select a winner for the match. In football, 30 additional minutes are used as the primary tiebreaker, and then penalty shootouts are used as a secondary tiebreaker.

II.1.e Tiebreaking Criteria (in a tournament)

In a tournament (or tournament stage) with points, tiebreaking criteria are used when two or more players have the same number of points to determine who qualifies for the next stage in the tournament (if there is one) or the final ranking between the tied players. In football, the goal difference is generally the first tiebreaker, with the total number of goals being the second tiebreaker. Since individual goals in matches are not simulated in this paper, players with an equal number of points in a tournament stage will be randomly ranked (i.e., ties will be broken randomly).

II.1.f Bye

A bye in a tournament is an automatic win generally given structurally by the tournament format when the number of players is not even or perfectly appropriate for the tournament format or sometimes given purposefully to give stronger players a better chance of winning.

II.2 Tournament Format/Structure

The way players will be matched against each other in the tournament, and how their results in those matches affect the final standings for the tournament.

II.2.a Bracket

A bracket style tournament format (also called knockout) refers to a tournament where losing is either eliminatory, or leads to a relegating bracket (which generally requires the player to win more matches).

II.2.b Single Elimination

The simplest bracket style tournament format. A total of 2^n players play in an n round bracket, with half of the players (the losing players in each match) being eliminated each round.

When the number of players is not 2^n, a certain number of players will be given a bye and immediately move on to the second match.

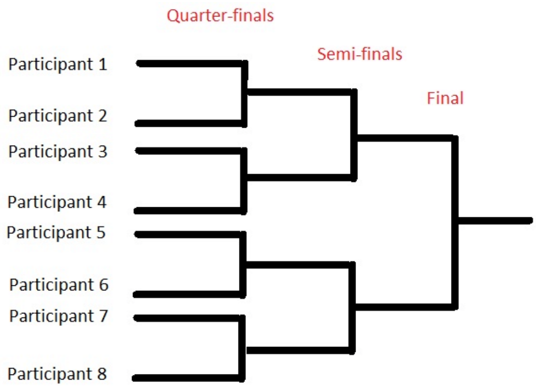

Figure 1: Single elimination tournament for 8 players

II.2.c Double/Triple/Multiple Elimination

A more complicated bracket style tournament format where players must lose two, three, or more times before being eliminated (instead of only once like in single elimination). Some players are given byes if the number of players does not equal 2^n. Some formats also feature players starting in the relegating bracket (often based on their seed) to purposefully give non-relegated players a much higher chance of winning.

II.2.d Points

Tournaments may have players face off against other players and earn points based on their result in each match. Players are then ranked by the number of points they have. The most common point systems are 2-1-0 (2 for a win, 1 for a tie, 0 for a loss; this system is identical to 1-0.5-0; this is the most prevalent system), 3-1-0 (rewards wins more, this is used by most football tournaments), and 3-2-1 (rewards playing more matches).

II.2.e Round Robin

A point-system tournament where all players face all other players once and gain points for their results. Double/triple/multiple round robin formats also exist, where each player faces off against every other player twice/thrice/multiple times.

II.2.f Swiss

A point-system tournament where players face off against other players with a similar number of points or record each round, while avoiding rematches (two players playing against each other more than once). Unlike other tournament formats where matches can easily be computed by hand, this format generally requires a computer and can involve solving a bipartite graph for the most optimal match pairings each round. A delayed Swiss refers to a Swiss tournament where two rounds are ready in advance, instead of only one, which allows players to know immediately what their next match will be without waiting for every other match in the round to finish.

II.2.g Multi-Stage Tournaments (Group Stage)

A majority of tournaments are comprised of stages, each essentially a “mini” tournament which are used to decide the seeding and/or the players who qualify for the next stage. Group stages refer to a tournament stage where players are placed into different groups, and each group plays its own “mini” tournament (generally, but not necessarily round robin). The (qualifying) players are then placed into the next stage of the tournament (which may be another group stage). Examples of common multi-stage tournament formats include a round robin group stage to a single elimination bracket stage; a Swiss stage followed by a bracket stage; or a Swiss stage, followed by a group stage, followed by a bracket stage.

II.2.h UEFA 2024/2025 Champions League

This new format (starting from the league phase, the qualification is not considered in this paper) is a two-stage tournament featuring 38 teams. In the first stage (the league phase), the teams are placed in four pots based on their seed, with the nine highest seed in the first pot, the nine next in the second, and so on. Each team is then randomly matched with two teams from each of the four pots (thus including their own) and play a total of eight matches. Teams are awarded 3 points for a win, 1 for a tie, 0 for a loss. The eight teams with the most points at the end of the league phase are immediately qualified for the round of 16 of the bracket stage. The next 16 teams are qualified for a playoff round to qualify into the round of 16 (a single match against another team). The last 12 players are eliminated. The winner of the bracket stage is the winner of the tournament. Each match in the league phase is comprised of one game. Matches in the bracket/knockout stage are two games long (except for the final which is a single game).

II.3 Seeding

The manner in which players, with a specific numbered seed going from 1 to the number of players, are placed in the tournament. Seeding is generally performed to ensure several different tournament design criteria. A higher seeded player is a player with a lower numbered seed (generally a better player).

II.3.a Random Seeding

The simplest seeding method, which involves randomly placing players in the tournament. It is a useful seeding method when attempting to avoid biases.

II.3.b Snake Seeding

This involves pairing the strongest players with the weakest players. For example, with 4 players, the pairings would be (1,4) and (2,3); 16 players would have (1,16), (2,15), (3,14), etc.

II.3.c Dutch Seeding

Players are paired in a “staircase” pattern, where the players are split into two halves, and then paired together. For example, with 4 players, the pairings would be (1,3) and (2,4); 16 players would have (1,9), (2,10), (3,11), etc.

II.3.d Monrad Seeding

Players are paired with the player immediately following (or preceding) them. The pairings would be (1,2), (3,4), (5,6), etc.

II.3.e Seeding in Bracket Style Tournaments

To place players in the bracket, the seeding method is done for each round, starting with the semi-finals, which are seeded according to the currently used seeding method, and then for every round before that by assuming the higher seed always wins the match. One could then obtain, for example, for a single elimination bracket with 8 players seeded with the snake seeding (the most common seeding for brackets) ([(1,8),(4,5)],[(2,7),(3,6)]); notice that, when assuming the higher seed wins, the bracket in the semi-finals would be [(1,4),(2,3)], which is the correct pairings for the snake seeding method for 4 players.

II.3.f Seeding in Swiss Tournaments

Seeding in the Swiss tournament is used in the first round (to determine initial pairings) and then, depending on the pairing rules, as the ultimate pairing rule if all other rules are insufficient or fail.

III. Efficacy Metric

Efficacy is a tournament design criterion which represents how likely the tournament is to give the correct ranking at the end of the tournament. If a tournament has three players, with seeds 1, 2, and

3, we expect the order of the players (the tournament ranking) to also be 1, 2, 3, in that order, with player 1 winning the tournament, and player 2 in second place.

There are a number of ways to compute the efficacy of a tournament. The simplest is to count the number of players who are ranked above a player with a higher seed (called inversion count). For example, a tournament with ranking 3, 2, 1 would have a count of 3, two from seed 3 being above 2 and 1, and one from seed 2 being above 1 (the closer the number to 0, the better). This already gives a very good approximate, but it fails to take into account other elements of a tournament.

We wish to give higher placements more importance, as people usually care more about the winners of a tournament than the eliminated players. One such method is from the literature (Sziklai, Biró, & Csató, 2022) which, for each player ranked at a higher position i than their seed j, adds a loss of

$$\sum_{k=i+1}^{j} -\frac{1}{\ln(k)}.$$

This is rather abstract but works very well in practice. In a three-player tournament, there are 6 possible outcomes. The inversion loss method matches the inversion count method except for 2, 3, 1; 3, 1, 2; and 3, 2, 1, where it gives the same loss to all of three rankings (as seen in table 1); it fails at distinguishing rankings where two (or more) players already placed in incorrect positions are further ranked incorrectly.

Table 1: The loss by Sziklai, Biró, & Csató (and Inversion Count) for each three player rankings

1st place	1	1	2	2	3	3
2nd	2	3	1	3	1	2
3rd	3	2	3	1	2	1
Penalty	0	-0.91	-1.44	-2.35	-2.35	-2.35
Count	0	-1	-1	-2	-2	-3

A much more intuitive way of thinking about loss inversion is through another method, which we will find is in fact equivalent, called prize loss. This method also introduces the prize distribution for the ranking into the method, making it very useful in discussing exactly which tournament formats are more efficacious (and, by extension, better). This method simply adds all the prize amount "lost" by players who would have deserved them. It is calculated by taking the expected prize a player would win and subtracting it from the actual prize obtained (with the prize loss fixed at 0 for a player if they have obtained more prize than expected). For example, in a four-player tournament with a prize distribution of 100,000 for first place, 50,000 for second place, 20,000 for third place and 0 for last place, if players place in order 3, 4, 1, 2, player 1 has lost 80,000 and player 2 has lost 50,000, therefore the total loss is 130,000 (the gains from players 3 and 4, equal to the total loss, are ignored).

We can see that by having a prize distribution of 0 for first place, $-\frac{1}{\ln(2)}$ for second place, $-\frac{1}{\ln(2)} - \frac{1}{\ln(3)}$ for third place, $-\frac{1}{\ln(2)} - \frac{1}{\ln(3)} - \frac{1}{\ln(4)}$ for fourth place, etc, we obtain the exact same total loss as the inversion loss method. Since prize loss is more intuitive and takes into account the prize distribution of the tournament directly, it will be the main method in which efficacy is calculated. However, it fails at differentiating some rankings from others because it values correct placements over correct ranking order. Whether one is more important than the other is a subjective matter, especially in tournaments with no prizes. In tournaments with prizes, we will consider prize misallocation (prize loss) to be paramount and therefore attempt to minimize it before considering correct ranking order.

Table 2: Efficacy of four player rankings, sorted by prize loss then by order penalty; prizes of, in order, 700, 300, 100 and 0

1st place	1	1	1	1	1	1	2	2	3	2	3	4	3	2	4	2	4	2	3	4	3	4	3	4
2nd	2	2	3	4	3	4	1	1	1	3	2	1	1	4	1	3	2	4	2	2	4	3	4	3
3rd	3	4	2	2	4	3	3	4	2	1	1	2	4	1	3	4	1	3	4	3	1	1	2	2
4th	4	3	4	3	2	2	4	3	4	4	4	3	2	3	2	1	3	1	1	1	2	2	1	1
Loss	0	-100	-200	-300	-300	-300	-400	-500	-600	-600	-600	-700	-700	-700	-700	-700	-700	-700	-700	-700	-900	-900	-900	-900
Inverse	0.00	-0.72	-0.91	-1.63	-1.63	-1.63	-1.44	-2.16	-2.35	-2.35	-2.35	-3.07	-3.07	-3.07	-3.07	-3.07	-3.07	-3.07	-3.07	-3.07	-3.98	-3.98	-3.98	-3.98
Order	0.00	-0.72	-0.91	-1.63	-1.82	-2.54	-1.44	-2.16	-2.35	-2.89	-3.80	-3.07	-3.26	-3.61	-3.98	-4.33	-4.52	-5.05	-5.24	-5.96	-4.71	-5.43	-6.15	-6.87
Count	0	-1	-1	-2	-2	-3	-1	-2	-2	-2	-3	-3	-3	-3	-4	-3	-4	-4	-4	-5	-4	-5	-5	-6
Loss+Order	0	-101	-201	-302	-302	-303	-401	-502	-602	-603	-604	-703	-703	-704	-704	-704	-705	-705	-705	-706	-905	-905	-906	-907

When prize loss is the same (or 0, because there is no prize pool), we propose here a novel method called order loss, which gives a weighted loss (just like inversion loss), for each player (with seed i) ranked above x different players with a higher seed, of

$$-\frac{x}{\ln(i+1)}.$$

(Table 2 presents all discussed methods for all possible four player rankings.) If order penalty is significantly smaller than the prize pool, summing the prize loss and order penalty gives a very useful metric to compare the efficacy of different rankings in a tournament.

However, given the use of Monte-Carlo simulations, we will use prize loss directly as the main efficacy metric here as the small results of order loss will be lost when summing with the prize loss; we are mainly interested in the average prize misallocation of the tournament formats.

IV. Fairness/Strategy-Proofness

Strategy-proofness generally refers to ensuring players will always attempt to win every single game without resolving to match-fixing or performing any other such strategies. Linked to this notion is fairness, which refers to ensuring that a stronger player always has a higher chance of winning the tournament, no matter how small that strength difference may be. This means that if the tournament is seeding players in a certain way, a higher seed must always have a higher chance of winning than any other lower seed. If this is not the case, players will attempt to purposefully lower their strength in order to guarantee a lower seed.

For fairness to be respected, for all integers k between one and the number of players in the tournament, the probability of a player finishing in the top k positions must at least equal the probability of any lower-seeded player finishing in the top k positions. This ensures that any player is always incentivized to get a better seed, as that would always improve their odds. More formally, this consists in checking that, given $P_k(i)$ the probability that seed i finishes in one of the top k positions,

$$P_k(i) \geq P_k(j), \forall i, j, k \in [1, 36], i < j.$$

Strategy-proofness is much more complicated and will not be explored in depth here as it is extremely dependent on the exact rules of the tournament. In summary, it requires checking, for every match, that losing does not improve that team's chances of finishing in the top k positions for any k. Using Monte-Carlo simulations to verify this would involve running nested Monte-Carlo simulations for every match which requires a different coding architecture altogether.

V. Methodology

V.1 Monte-Carlo Simulations

Monte-Carlo simulations will be used to determine both the efficacy and fairness of the investigated formats. This approach is preferable in these cases as a purely deterministic or brute-force approach (calculating the exact probabilities of every single actual and possible match) is too computationally expensive for multi-stage tournaments (it is however feasible for single stage bracket or round robin tournaments). Monte-Carlo simulations offer a very accurate approximation of the actual results in a reasonable amount of time.

V.2 Bradley-Terry Model

The Bradley-Terry model is a pairwise probability model used to determine the probability of an event occurring (such as "team x winning the game against team y") based on prior scores given to the two items (such as the strengths of team x and y).

According to the Bradley-Terry model, when p_x and p_y are the strengths of two players/teams in a game/sport, in the range $(0, \infty)$, the probability that team x beats team y in a game is

$$\frac{p_x}{p_x + p_y}.$$

The Bradley-Terry model does not take into account ties. Since football has a prevalence of ties, ignoring them is undesired. To model ties, we introduce t such that

$$\frac{t}{p_x + p_y + t}$$

is the probability that a tie occurs, and we modify the win probability of team x to

$$\frac{p_x}{p_x + p_y + t}.$$

We want to find t such that

$$\frac{t}{p_x + p_y + t} = g(p_x, p_y, b)$$

where, for some given $b \in (0, 1]$

$$g(p_x, p_x, b) = b \ \forall p_x \in (0, \infty),$$

$$g(p_x, p_y, b) < b \text{ when } p_x \neq p_y \text{ and}$$

$$g(p_x, p_y, b) = g(p_y, p_x, b) \ \forall p_x, p_y \in (0, \infty).$$

Intuitively, we want g to be the probability of a tie between players with strength p_x and p_y, with a maximum probability of b achieved when $x = y$.

We find one possible such function g (when restricting x and y to $[0, 1]$, but changing the coefficient π to some other coefficient would allow for more values) with

$$g(x, y, b) = b\,\frac{1 + \cos(\pi(x - y))}{2}.$$

We can now use this to find t. We have

$$\frac{t}{x + y + t} = g(x, y, b) => t = \frac{g(x, y, b)(x + y)}{1 - g(x, y, b)}$$

and so

$$t = \frac{b\frac{1+\cos\left(\pi(x-y)\right)}{2}}{1-b\frac{1+\cos\left(\pi(x-y)\right)}{2}}(x+y)$$

which is quite complex but works rather well in practice. Other functions g might be more appropriate depending on the restriction on x and y.

The value of b (referred to later as the tie coefficient) will be set at different values, namely 0.5, 0.3, 0.1, and 0.

V.3 Team Strengths Distribution

When performing the simulations, the strengths of the players/teams need to be chosen. While the exact numbers are relatively unimportant, the overall distribution of the ratings is.

To accommodate the restriction of the function g from the previous section, we will divide all strengths by the highest strength of all teams. Strengths also still need to be positive. This leaves us with a distribution in $(0, 1]$.

Therefore, we wish to find distributions for the simulations inside $(0, 1]$. We require that teams with a higher seed have, on average, the same or a higher strength than lower seeds. If this is not the case the results are inapplicable. It is also preferable that teams with a higher seed always have a higher strength, but this is not always the case in actual tournaments. These requirements also make sense in that the seeding for the tournament will also generally follow these criteria (a higher seeded player will have a higher strength than a lower seeded player).

Therefore, the simplest distribution is an ordered uniform distribution in $(0, 1]$, with the highest strength given to the player with seed 1, the second highest to seed 2, etc., which will allow us to use this distribution both for efficacy and fairness.

Instead of ordering the distribution, the other possibility is a random sampling which is then divided by x^{i-1} where x is a chosen variable (such as 2, 10, or 100) and i is the player's seed. Therefore, for example, with $x = 2$ and sampling from a continuous distribution in $(0, 1)$, we have seed 1 with a strength from a continuous random variable in $(0, 1)$, seed 2 in $(0, 0.5)$, seed 3 in $(0, 0.25)$, and so on. This skews the overall distribution of the strengths towards 0 which should generally allow the tournament formats to give a very accurate ranking. While more prize loss could potentially be expected with such a distribution (since players are sometimes given a strength which is higher than other players with a higher seed), the strength for players with very low seeds will be taken in intervals such as $(0, 10^{-11})$ which means high seeds will almost always certainly win (such an assured victory is almost never the case in real life).

Finally, using data from previous matches between the 36 teams in the Champion's League, two more sets of ordered strength distributions were generated. One using only the betting odds from betting venues for either one team's or another team's win, calculated by iterating over the odds until each team's strength matched the odds as closely as possible, and the other using Glicko, a popular rating algorithm which can give each team a rating based on their past results (Appendix A presents their distribution).

Different prize losses for different distributions are acceptable and expected, as the objective is to compare different tournament formats, not the different distribution of strengths. As long as the same distributions are applied to all simulations, we can compare such prize losses for each distribution effectively.

V.4 Tournament Formats Studied

To fairly compare the new Champion's League format with other tournaments, we need teams to play the same number of matches. We study here two main alternative candidates for a total of three tournament formats studied.

All three formats are two-stage 36-team tournaments starting with a point-based format where a win gives 3 points, a tie 1 point, and a loss 0, with 24 teams qualifying for the bracket stage and the top 8 of them getting a bye to the round of 16 immediately.

V.4.i Bracket Stage (Final Stage for all formats)

The bracket stage remains the same between all three formats. The top 8 teams in the first stage of the tournament are qualified for the round of 16, and teams ranking 9th to 24th are placed in a playoff round before qualifying to the round of 16. All matches, except for the final, are two games long, the final being a single game.

The bracket stage features two-game matches, and therefore the probabilities from the Bradley-Terry model are used slightly differently than for a single-game match. When simulating a two-game match and since individual goals are not simulated, two games are simulated for the match: if there is a tie and a win for a team, or two wins for a team, that team wins the match, otherwise a third game is simulated with the tie probability increased by 50% (to simulate a 30 mins overtime); if this third game is also a tie, the team with a higher strength is given a 55% chance of winning, and the other team is given 45% chance (to simulate penalty shootouts). Since the bracket stage is the same between all formats, the exact algorithm of a two-legged match will not significantly alter results between the tournament formats and, therefore, this approximation was chosen.

V.4.ii League Phase (First Stage)

The League Phase of the UEFA Champion's League for 2024/2025 is the second to last stage of the tournament before the bracket stage. It features 36 teams placed into four pots based on their seed. The 9 highest seeds in the first pot, the 9 next in the second pot, and so on. Each team is then matched with two teams from each pot (including their own pot) to play a single game, for a total of 8 games/matches. All players are then placed in the same leaderboard for qualification to the final stage.

V.4.iii Group Round Robin (Alternative First Stage)

A simple alternative to the League Phase is a group round robin stage. In this case we have 4 groups of 9 teams playing one game against all teams in their own group. Leaderboards are only group specific, with the top 2 teams in each group qualifying for the round of 16 immediately (for a total of 8 teams), and the 3rd to 6th teams in each group qualifying for playoffs (for a total of 24 teams qualifying for the bracket stage).

V.4.iv Swiss Stage (Alternative First Stage)

The third alternative studied will be a Swiss stage. Teams will be matched round per round based on their current number of points (instead of scheduling all eight matches at the beginning of the stage). A delayed Swiss may also be preferable (where two rounds are prepared in advance instead of one) for practicality purposes so that teams may already know where their next match is after finishing the previous one. The Swiss stage groups all players in a single leaderboard for qualification to the final stage and qualifies players in the same way as the League Phase.

It is important to note that all first stages here have the exact same number of matches. All teams will play 8 matches in the first stage, yielding the same amount of information regarding the "true" ranking of the teams.

V.5 Swiss Algorithm and Tournament

The Swiss tournament, being relatively more complex than others, will be elaborated upon.

The two most general rules followed by all Swiss algorithms are that rematches are forbidden (players may never be paired with one another more than once) and players with a similar score are paired together. There are a number of ways achieve this and the exact pairing algorithm can differ. One of the most complex Swiss algorithms is in chess, which is impractical to do by hand and requires a program to make the pairings. FIDE (the International Chess Federation, "Fédération Internationale Des Échecs" in French) requires that player never play white or black significantly more than the other (no more than a difference of +2 or -2) as well as requiring that players never play the same colour three times in a row. (FIDE, 2024) These restrictions disallow some pairings between players who, for example, both have a difference of +2 or have played white in the last two games, even if they have never played against each other and have the same number of points. Similar restrictions could be put in place for football tournaments with "home" and "away" games instead of white and black.

Major online tournament websites generally follow a simpler variation of the FIDE Swiss algorithm, with significantly less restrictions. These place players into groups based on the current number of points they have where players with the same number of points are grouped together. Groups with an odd number of players drop one player "down" to the next group (which will ensure all groups have an even number of players). The algorithm then attempts to pair all players with another player in the same group as them (a pairing is "allowed" if those two players have never been paired together). If all pairings are allowed, the algorithm is done for this round. If one group cannot be paired by any means, it is merged with the next group down (or up if it is the last group) and pairings are attempted again until valid pairings are found. This algorithm has been implemented for Monte-Carlo simulations but is terribly inefficient when running hundreds of simulations, therefore the total number of simulations for this algorithm will be kept low.

Another much more efficient algorithm used for other Monte-Carlo simulations of tournaments in the literature (Sziklai, Biró, & Csató, 2022) instead pairs players directly, without making any groups based on score. All players are ordered based on their points (ties are broken with the Buchholz system, which simply sums the number of points of all encountered opponents, and then broken by seed) and, starting with the first player, paired with the next best player which they have not yet played against. If a player cannot be matched (because that player has already played against all other unmatched players), the previous pairing is undone and the previously paired player is now matched with the next highest player they can be matched with (and if this player cannot be matched again, the last pairing is also undone, which may be repeated ad infinitum, until all players have been paired). This yields a very efficient algorithm but which, by construction, may not necessarily pair the players with the worst score in the tournament most effectively.

For our purposes, we will call the latter algorithm the alternate Swiss Tournament stage.

V.6 Efficacy: Prize Loss

To calculate prize loss, the tournament requires a prize pool and rules for distributing that prize pool. Since we are comparing the Champion's League with other tournaments, we will use the Champion's League prize pool and distribution rules to calculate prize loss. The "performance related" prize money is distributed as follows: (UEFA, 2024)

- During the League Phase (the first stage), teams receive bonuses:

- Performances bonuses in amount of €700k per point obtained. This will be ignored when calculating prize loss as it is not distributed according to the final or intermediate ranking of a tournament but rather as an incentive to win individual matches and avoid draws.
- League Ranking Bonuses in amount of €275k per share with 666 total shares. Teams will receive a number of shares equal to 37 minus their position (36 shares for first place, 35 for second, 34 for third, and so on, with a single share for 36th place).
- A bonus of €1M for clubs ranked 16th or above at the end of the League Phase.
- A bonus of €1M for clubs ranked 8th or above at the end of the League Phase (for a cumulative total bonus of €2M).

- During the bracket/knockout stage, teams will receive a flat amount of:
  - €1M for teams eliminated in the playoff round,
  - €12M for teams eliminated in the round of 16,
  - €24.5M for teams eliminated in the quarter-finals,
  - €39.5M for teams eliminated in the semi-finals,
  - €58M for the runner-up team eliminated in the final,
  - €64.5M for the winner of the tournament plus an additional €4M for participation in the European Super Cup (a single game between the winners of the Champion's League and Europa League) with an additional €1M bonus for winning that match (for the purposes of calculating prize loss, the prize for first place will be aggregated to €69.5M).

These prize distribution rules (with the above addendums for first place and individual match performance bonuses) will be used directly in both the Champion's League format simulations (League Phase followed by Bracket Stage) and the Swiss simulations (Swiss Stage followed by Bracket Stage). For the round robin group stage format, an additional modification is made as to how shares are distributed, with teams winning 34.5 shares for first place in their group, 30.5 shares for second place, 26.5 shares for third place, and so on, with 2.5 shares for the teams in ninth place in each of the four groups, for the same total of 666 shares.

VI. Efficacy Testing

We first run simulations to test efficacy. Again, we have three different tournament formats, the third of which has two different pairing algorithms. We also test each with different strength distributions and with different tie coefficient values.

All of the results are based on 100 000 simulations for each combination of distribution and tie coefficient, except for the Swiss format which only has 200 (the alternative Swiss format still has 100 000 simulations).

Thanks to previous research in the field (Devriesere, Csató, & Goossens, 2024) (Sziklai, Biró, & Csató, 2022), and since all formats have the exact same number of matches, we expect to see the Swiss tournament perform slightly if not significantly better than the other two formats, with the non-alternative Swiss performing slightly better than the alternative Swiss (due to its more efficacious algorithm). We also expect to see the group stage perform better than the league phase due to its full and balanced round robin structure (all teams play against all other teams in their own group).

All results are visible in Figure 2, with the exact prize loss of each Monte-Carlo run in Appendix B.

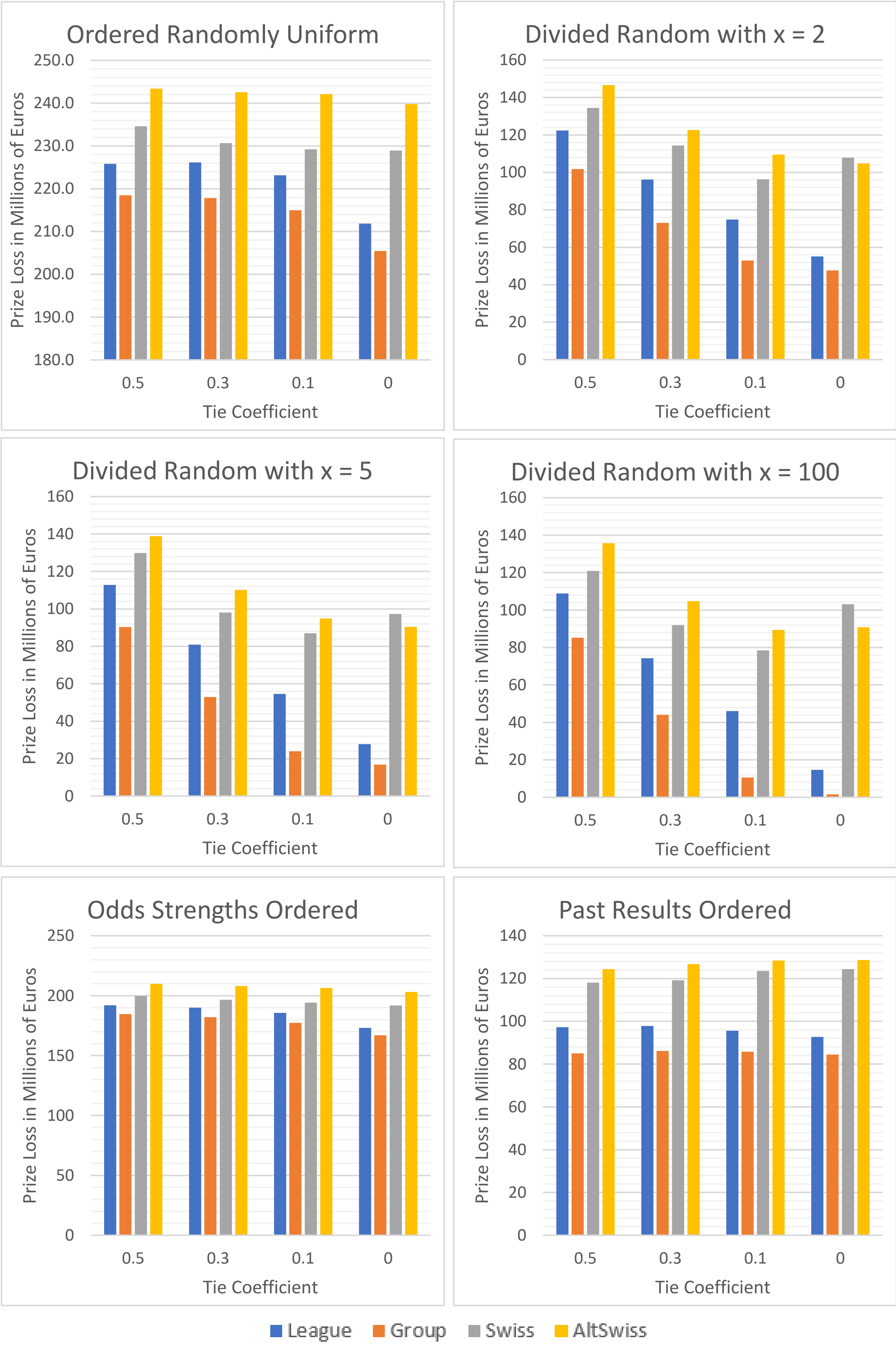

Figure 2: The simulated prize loss for six different strength distributions and four tie coefficients each.

The trend is perfectly clear in all simulations. The group stage always performs better than the league phase, which itself always performs better than the Swiss stage. This is contrary to our hypothesis. While the group stage was expected to perform better and has indeed done so, the Swiss stage was also expected to perform better than that, and it instead shows more prize loss than the other formats, including the league phase. This discrepancy will be studied more in depth in an attempt to understand whether this remains true when changing basic assumptions on the tournament and players strength distribution as well as the metric for efficacy.

Some other interesting or unexpected results include the higher prize loss in some cases when the tie coefficient is reduced. While reducing the tie coefficient usually also predictably reduces prize loss (since reducing ties might allow stronger players to win points more consistently), some tournament formats (most specifically the non-alternative Swiss) paired with some specific distributions (especially ones where the strengths are far apart but linear on $(0,1)$), seem to benefit unreasonably well from having ties. This is likely caused by how the pairing algorithm in the non-alternative Swiss behaves in the latter rounds (where score groups may easily be merged with each other due to the no-rematch pairing rule), which causes some lower strength teams to steal a full 3 points from strong teams instead of a much more likely 1 point tie for both teams if the tie coefficient were higher. This would also explain why the prize loss is sometimes higher for the non-alternative Swiss compared to the alternative Swiss, since the alternative Swiss pairing algorithm makes no such groups based on score and simply attempts to pair the best players together.

But this does not explain why other tournament formats also sometimes benefit from allowing some ties. This is likely caused by the tiebreaker rules in the bracket stage, where allowing ties (tie coefficient greater than 0) would mean stronger teams now have a chance of playing an additional 30 mins if tied, whereas if the tie coefficient were 0, that would have been a loss. Since stronger teams always want to play more to have more chances of proving themselves, disallowing ties restricts their ability to do so.

VII. Swiss Lower Efficacy Understanding

It was common knowledge and then verified through simulations (Sziklai, Biró, & Csató, 2022) that the Swiss tournament is the most efficacious tournament against other tournaments with the same number of matches (compared to standard tournament formats, not necessarily the most efficacious against all possible tournament formats). However, our simulations show the opposite, where a two-stage tournament starting with a group stage or even the Champion's League "League Phase" is more efficacious than a two-stage tournament starting with a Swiss stage. We now attempt to make the Swiss tournament more efficacious than other tournaments by changing the Monte-Carlo simulation's parameters and code.

We first wish to see if choosing the specific prizes from the UEFA's Champion's League may have unfairly affected the Swiss tournament as the Swiss tournament is supposedly better than other tournament formats at getting the overall ranking of all players rather than only the first-place player. However, whether choosing prize loss, inverse loss, order loss, or even inverse count, the two-stage Swiss tournament still performs worse on efficacy.

To eliminate some more variables, we will now only calculate the efficacy metric on the rankings after the first stage, so we perform simulations on a single stage tournament and do not take into account the bracket stage. With the results in figure 3, this still does not manage to make the Swiss tournament more efficacious compared to the other formats (the league phase on its own and the nine-player group stage round robin four times, for 36 players).

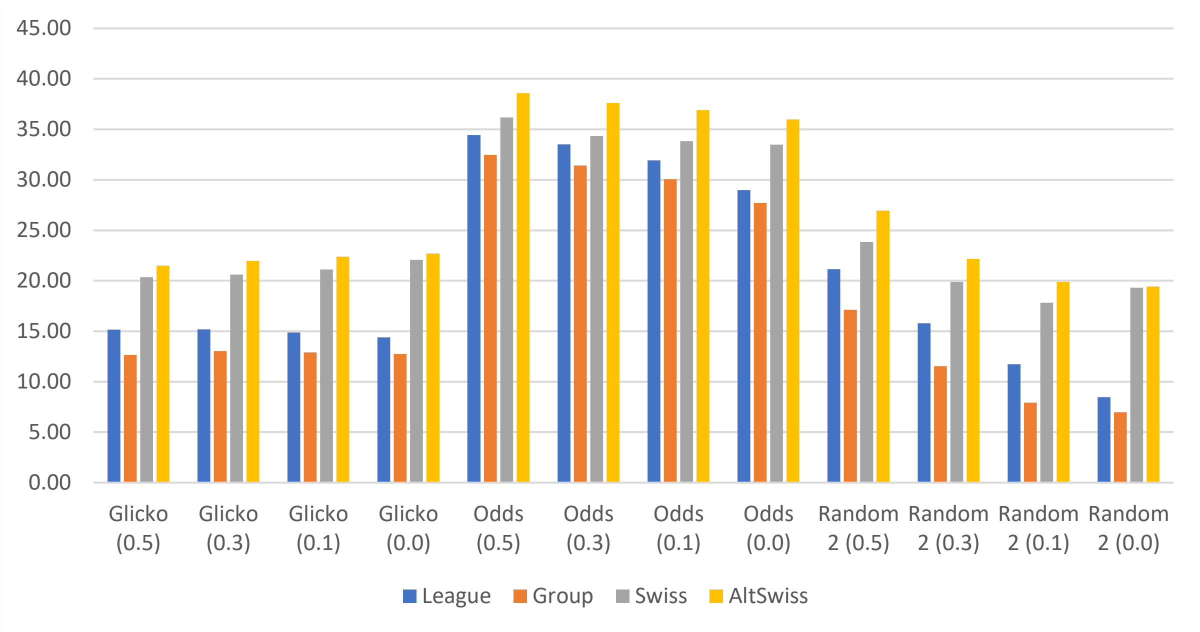

Figure 3: Inverse Loss for Different Strength Distributions (with Different Tie Coefficients) of only the first stage

The football format also rewards wins more than draws, giving three points for a win rather than the usual two. However, changing this to two points actually makes the tournament less efficacious (as long as the tie coefficient does not equal 0 of course, at which point giving three or two points for a win is equivalent), so this does not seem to be the root cause of the efficacy problem.

At this point, apart from a bug in the code causing the Swiss tournament to have less efficacy (in one way or another), there remains only one major assumption (when testing a 36 player, 8 round tournament), which is the Bradley-Terry model. We test another model where the probability of player A winning against player B is given by

$$P_{AB} = 0.5 + \text{skill } \times \big(\text{seed(B)} - \text{seed(A)}\big)/100$$

where "skill" is a parameter. (Sziklai, Biró, & Csató, 2022) Unlike the Bradley-Terry model, this probability model is linear and, in fact, the probability of a player winning a match could exceed 1 according to the formula if the seeds are far enough apart and the skill parameter set high enough, at which point the win is guaranteed (the probability is clamped between $[0, 1]$).

We test this new model with skill set to 1, 5, and 10. Even then, the inverse loss is still higher for the Swiss tournament compared to the group stage and the league phase (as seen in figure 4).

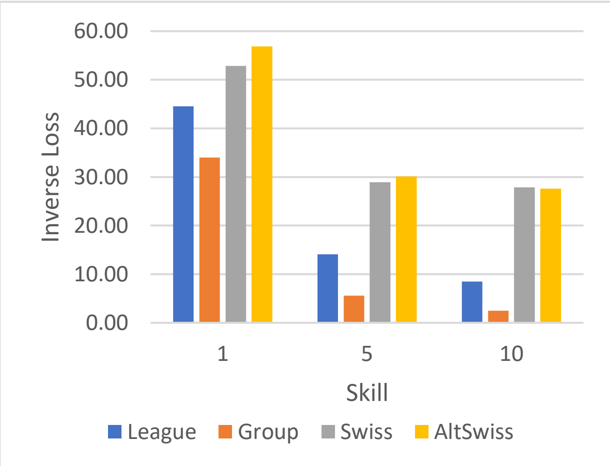

Figure 4: Inverse Loss for different skill values for the first stage

Given the conflict in results (with (Sziklai, Biró, & Csató, 2022)), it remains inconclusive as to whether Swiss is more or less efficacious than other tournaments given the same number of matches. Our simulations here were only on the specific case of 36 players/teams with 8 rounds (and so 8 matches per player) for a total of 144 matches (for all tournament formats). While simulations on the round robin group stage when only the first stage is being investigated may not be particularly applicable (ending the tournament after a group stage would not yield full rankings), the results for the league phase and swiss stage are comparable as they both give full player rankings after 8 rounds, and the Swiss stage was consistently less efficacious than the league phase.

VIII. Fairness Testing

We now test the UEFA Champion's League for fairness. This consists of checking that, given $P_k(i)$ the probability that seed i finishes in one of the top k positions,

$$P_k(i) \geq P_k(j), \forall i, j, k \in [1, 36], i < j.$$

This essentially means that for a tournament to respect the fairness criteria, a higher seed always needs to have a higher probability of placing in the top k positions compared to a lower seed, for any k, as long as some strength requirements are respected, namely that the strength of a higher seeded player is equal or higher than that of a lower seeded player, and this must be true regardless of the strength distribution.

Given that a bracket stage does not distinguish between players eliminated during the same round, we only check the above for k equal to 1, 2, 4, 8, 16, and 24 (we assume results for positions 25 to 35, teams eliminated in the league phase, are unimportant enough that teams would not purposefully fix matches to gain an advantage for these lower positions).

Results can be seen in figure 5 with 100 000 simulations for each tested combination of distribution and tie coefficient. Practically all results do not break fairness, a couple of Monte-Carlo simulations break fairness by a miniscule margin (less than ten in a hundred thousand simulations) and are therefore ignored; however, with some specific combinations of tie coefficients and strength distributions, we see a slightly higher percentage for seeds 10, 19, and 28 compared to seeds 9, 18, and 27 respectively.

Table 3: Probabilities of Different Seeds Achieving Different Positions for Modified Odds Calculated Strengths with the Tie Coefficient Set to 0.5

	1st	Top 2	Top 4	Top 8	Top 16	Top 24
Seed 9	2.30%	5.87%	14.53%	33.42%	68.21%	90.53%
Seed 10	2.33%	5.92%	14.65%	33.63%	68.45%	90.81%
Seed 18	0.30%	1.07%	4.00%	13.23%	40.56%	73.22%
Seed 19	0.30%	1.07%	3.97%	13.23%	40.75%	73.46%
Seed 27	0.01%	0.08%	0.54%	3.08%	16.17%	44.34%
Seed 28	0.01%	0.08%	0.54%	3.11%	16.34%	44.82%

To ensure this is consistent, we can take a strength distribution and copy the strengths of seed 9, 18, and 27 to those of seeds 10, 19, and 28 respectively (this still respects fairness requirements for strength distribution because the strength is not lower than that of lower seeds); we will take the odds calculated strengths with a tie coefficient of 0.5, since they seem to present the most discrepancy for those seeds. If fairness were to be respected, we should see equal, or higher, probabilities of seeds 9, 18, and 27 placing higher than those of seeds 10, 19, and 28 respectively. To reduce randomness further, we run a million simulations, with the results visible in table 3. Lower seeds are consistently placing higher on average for certain top positions, most notably for seed 10 who is placing higher up to one percent more often (between 0.03 and 0.3 percentage points) compared to seed 9.

This can be explained by how matches are assigned during the league phase. During the league phase, teams are assigned two other teams from each hat, which gives them 9 possible teams for each hat, except their own where it is only 8. This discrepancy means that a team with, for example, seed 9 will play two matches against teams with seed 1-8 and two matches against teams with seed 10-18, while the team with seed 10 will play two matches against teams with seed 1-9 and two

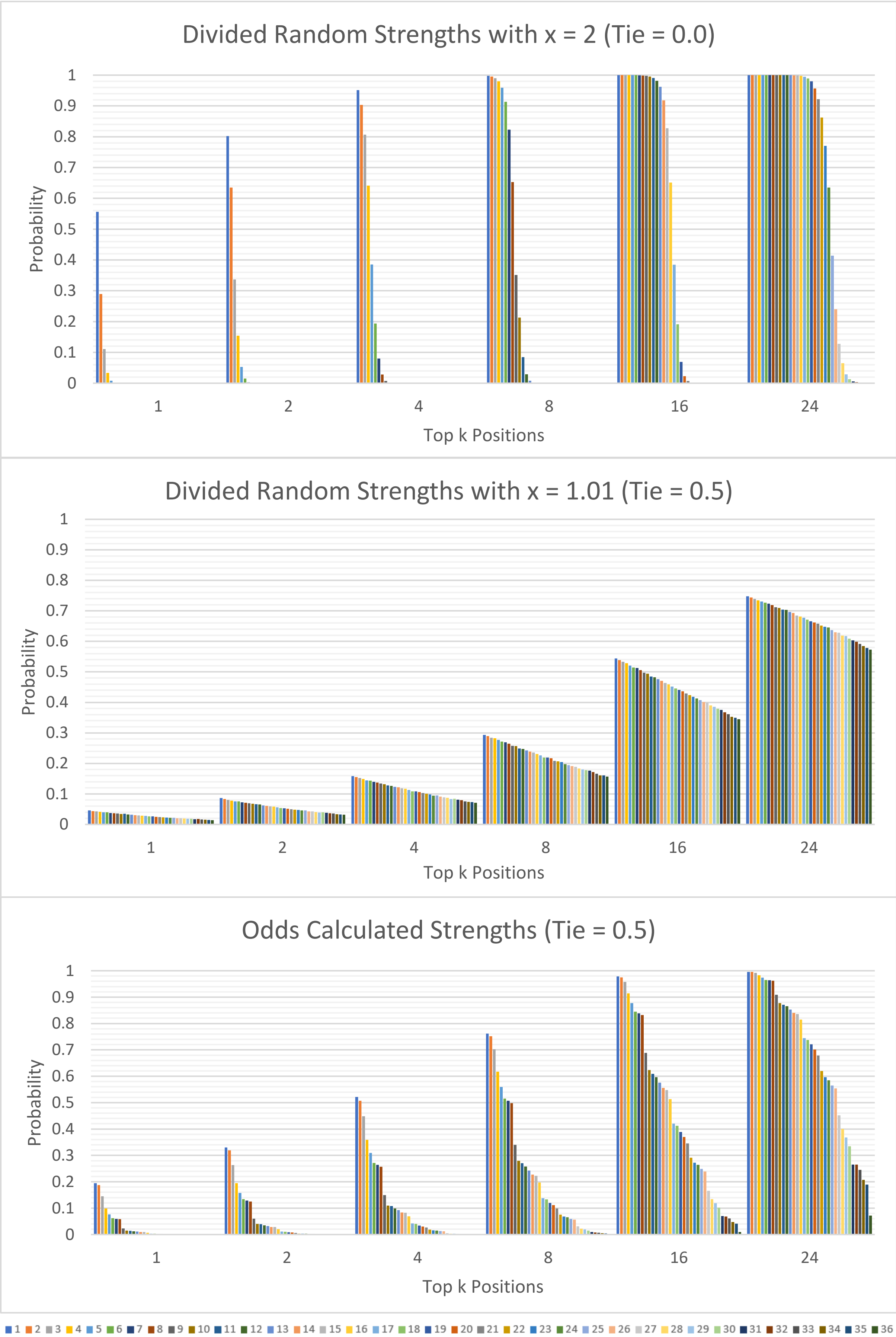

Figure 5: Some Monte-Carlo Simulations Run to Test Fairness with Different Distributions and Tie Coefficients, Each Bar Represents the Probability of Seed i Being in the Top k, 100 000 Simulations in Each Run

matches against teams with seed 11-18. This match arrangement leaves seed 10 with a $\frac{2}{9}$ chance of playing against seed 9, which would mean seed 10 is playing against the "easiest" team in the first hat, avoiding a potentially harder match, while seed 9 is forced to play against the strongest team in the second hat, missing an opportunity for an easier match against seeds 11 to 18.

Since some higher seeds do have a lower chance of winning compared to some lower seeds, the UEFA Champion's League does not respect fairness. Some simple changes to the league phase would fix this, either by reducing the number of matches a team plays in their own hat by 1, so they would be matched with two teams from all other hats, but only one from their own hat (but this would reduce the total number of matches played by all teams to seven); or by reducing the number of hats to three and then each team is matched with three other teams from each hat, except their own where they are matched with only two teams (this would conserve a total of eight matches for each team). Either way, reducing the number of matches by one in the team's own hat would fix the fairness issue for this tournament format. This works because a team getting placed into a higher hat would reward them with one less match against those higher skilled teams. This would essentially be a small "reward" for higher seeded teams as they will face slightly easier teams and would make it slightly harder for teams in lower hats.

IX. Conclusion

The UEFA lists three broad goals for the new Champion's League format, but one of them in particular states that "the new format will introduce a better competitive balance between all the teams, with the possibility for each team to play opponents of a similar competitive level throughout the league phase." (UEFA, 2024) When comparing a round robin group stage with the current league phase, it would be strange to call this goal fulfilled when the group stage already fulfills this goal while being more efficacious and respecting fairness.

The change in the tournament format is an odd choice from the point of view of efficacy and fairness, as better alternatives are available, namely the standard round robin group stage followed by a bracket stage. But the UEFA Champion's League is reasonably efficacious (being only slightly worse than the round robin group stage) and while fairness is not respected, the lower seeds are gaining less than 1% more win probability, which is practically negligeable. Nevertheless, a very small fix to the league phase would eliminate this issue and allow the tournament to respect fairness.

There remains the issue of the Swiss tournament format, which in this paper was less efficacious than the other two compared tournament formats, while having been deemed more efficacious in other research. This inconsistency should be more closely investigated. The chosen algorithm for Swiss pairings also remains an open area of research. Other tournament formats could also be considered, but these two other tournament formats were the closest to Champion's League format. Additionally, this paper focused on the Bradley-Terry model, but other models exist, and even a model allowing any sequence of pairwise comparisons could be considered when testing efficacy. Finally, more research to test the importance of given matches, especially in regards to the Champion's League, and allow for a metric of importance for a tournament as a whole could allow tournament designers to more carefully select the tournament format with stakeholders' goals in mind.

Bibliography

Devriesere, K., Csató, L., & Goossens, D. (2024). Tournament design: A review from an operational research perspective. *arXiv*.

FIDE. (2024, September). *FIDE Handbook*. Retrieved from International Chess Federation: https://handbook.fide.com/

ResearchGate. (2025, January 21). *An Analysis of Tournament Structure - Scientific Figure*. Retrieved from ResearchGate: https://www.researchgate.net/figure/Single-elimination-tournament-for-8-participants_fig1_310953167

Schilling, M. F. (1994). The Importance of a Game. *Mathematics Magazine*.

Sziklai, B. R., Biró, P., & Csató, L. (2022). The efficacy of tournament designs. *arXiv*.

UEFA. (2024, March 22). *Distributions to Clubs of UEFA*. Retrieved from UEFA: https://editorial.uefa.com/resources/028b-1a7880138a24-7a993e2e33d1-1000/20240322_circular_2024_13_en.pdf

UEFA. (2024, June 12). *New format for Champions League post-2024*. Retrieved from UEFA: Champion's League: https://www.uefa.com/uefachampionsleague/news/0268-12157d69ce2d-9f011c70f6fa-1000--new-format-for-champions-league-post-2024-everything-you-ne/

Special thanks to Vianney Perchet and Rémi Verronneau for their help in writing this paper.

Appendix A: Odds and Past Match Calculated Strengths

Table 5: Odds Calculated Strengths for the 36 teams in the Champion's League

Bayern Munich	0.998191004	AC Milan	0.413334496	Stuttgart	0.266408504	Salzburg	0.130907962
Manchester City	0.982206499	Dortmund	0.406881392	PSV	0.257114055	Brest	0.122632812
Liverpool	0.885344367	Atl. Madrid	0.396849851	Club Brugge KV	0.243354157	Sparta Prague	0.113906213
Arsenal	0.763226061	Atalanta	0.383376095	Celtic	0.212975009	Sturm Graz	0.094957083
Real Madrid	0.694297512	Sporting CP	0.371841891	Monaco	0.205293421	Shakhtar Donetsk	0.094872139
Barcelona	0.645831185	Aston Villa	0.368923225	Feyenoord	0.201937636	Crvena zvezda	0.089986984
Bayer Leverkusen	0.641704231	Juventus	0.34538772	Girona	0.195340044	D. Zagreb	0.079798211
PSG	0.637695708	Benfica	0.28294984	Lille	0.191311982	Young Boys	0.075826308
Inter	0.474135371	RB Leipzig	0.280422385	Bologna	0.150376555	Slovan Bratislava	0.040225964

Table 4: Past Matches Calculated Strengths for the 36 Teams in the Champion's League

Liverpool	2272.750273	Bayern Munich	1777.385332	Feyenoord	1603.821237	Shakhtar Donetsk	1099.929506
Barcelona	2062.309897	Dortmund	1768.41199	Monaco	1596.692422	Bologna	1030.721207
Arsenal	1958.161083	AC Milan	1745.96265	Real Madrid	1558.55052	Crvena zvezda	1004.457113
Bayer Leverkusen	1947.564511	Atl. Madrid	1734.181535	PSV	1526.190719	Sturm Graz	991.5765846
Brest	1924.393216	Atalanta	1710.535947	Manchester City	1515.872898	Salzburg	972.0146563
Lille	1882.411726	Club Brugge KV	1670.693443	D. Zagreb	1488.469176	Girona	951.0621497
Inter	1877.577605	Celtic	1632.160609	Stuttgart	1340.124818	RB Leipzig	768.9005696
Aston Villa	1856.882636	Sporting CP	1621.352731	PSG	1334.036741	Slovan Bratislava	700.3979585
Juventus	1777.977266	Benfica	1614.390542	Sparta Prague	1109.585632	Young Boys	691.1971979

Appendix B: Exact Efficacy Simulation Results

		Loss in Euros of Tournament Formats			
Tie	Strength Distribution	League	Group	Swiss	AltSwiss
0.5	Ordered Random	225798669.5	218497690.0	234595875.0	243389985.0
0.3	Ordered Random	226130288.3	217825971.0	230674750.0	242568438.5

0.1	Ordered Random	223164676.5	214961145.0	229207625.0	242130349.3
0	Ordered Random	211882908.0	205436376.0	228927125.0	239782111.8
0.5	Random x = 2	122370981.3	101845655	134499625	146686100.8
0.3	Random x = 2	96239650.5	73052490	114364625	122683739.5
0.1	Random x = 2	74845866	52943692	96343875	109480414
0	Random x = 2	55088163	47632170	107892875	104837945
0.5	Random x = 5	112824871.8	90258105	129830000	138830727.8
0.3	Random x = 5	80919294.25	52884467	98037000	110113652.8
0.1	Random x = 5	54536186.75	23876294	87010500	94883250.75
0	Random x = 5	27697156	16807652	97285875	90489439.5
0.5	Random x = 100	108876583.5	85190275	120915875	135657940.8
0.3	Random x = 100	74329755.25	43986686	91967000	104859447.5
0.1	Random x = 100	46023161.25	10533892	78464875	89477544.25
0	Random x = 100	14660182	1589287	103206250	90867840.25
0.5	Odds	191939923.8	184667097	199711625	209884419.5
0.3	Odds	189930203	182077600	196547875	208097457.5
0.1	Odds	185547430	177275568	194099375	206564722.8
0	Odds	173142876.8	166901418	191777250	203026934.3
0.5	Past Matches	97312510.5	85066530	118030500	124437963.5
0.3	Past Matches	97810117.25	86143739	119173250	126692676
0.1	Past Matches	95626474	85838810	123527750	128387241
0	Past Matches	92654811	84484435	124436250	128594378.5