

# Modern Playoff Tournament-Design Analysis across Sports: A Bayesian Modeling and Advanced Simulation Study

Hassan Rafique and Shane Sanders

Department of Sport Analytics, Falk College of Sport, Syracuse University

## 1. Introduction

The global spectator sports market was valued at \$178.74 billion in 2024, following robust annual growth in recent years (Yahoo Finance, 2025). Buoyed by the digital age and recent competition format experimentation, e.g., in the Summer Olympics, Major League Baseball, Indian Premier League (IPL) Cricket, Pakistan Super League (PSL) Cricket, the FIFA World Cup, the NBA Playoffs, and College Football, sports viewership increases have supported these revenue gains. The 2020 Summer Olympics attracted approximately 3 billion viewers, the World Series approximately 238 million total unique viewers, the IPL final more than 1 billion digital and television viewers, the 2022 FIFA World Cup Final approximately 1.5 billion viewers, and the 2025 Super Bowl a record 127.7 million U.S. viewers.

Rules are fundamental to sport's functioning. Moreover, the sport business and sport management literatures have identified tournament and competition design, an increasingly important function of sport operations and analytics groups within leagues, as a vital aspect of continued growth in sport attendance [1,2]. Tournament design refers to the set of rules governing tournament structure, encompassing the group of contestants, competitive format, schedule, and ranking [3]. Thus, "designing an optimal contest is both a matter of significant financial concern for the organizers, participating individuals, and teams, and a matter of consuming personal interest for millions of fans" (Szymanski 2003)[4].

Ostensibly, this level of importance helps to explain several recent, high-profile cases of tournament design experimentation. Even before such experimentation, however, new tournament designs can be optimized vis-à-vis key, established axioms of tournament viability. Within such an approach, axioms are specified from the literature and designs pre-tested using simulation, advanced modeling, and optimization approaches. This model-based approach is much less opportunity costly than the approach of real-world trial and error. For example, the College Football Playoff (CFP) has undertaken several modifications to varying success in recent years. Some of these stepping-stone tournament design iterations may have been circumvented through the proactive use of a design approach that optimizes key desired tournament features. In the traditional tournament design literature, these axiomatic features include: 1) **Efficacy**: revelation of true team-ranking, in expectation, via a play of the tournament; 2) **Fairness**: Team win likelihoods increase in true (perhaps unknown but to the modeler) team strengths; 3) **Attractiveness**: The tournament generates excitement, exposure, and financial viability to the League.

Herein, we evaluate newly-adopted hybrid tournament designs from IPL, PSL, and BPL Cricket, as well as the NBA Conference-level Play-In Tournaments and the current NCAA CFP, against classical

designs from post-group play in certain Olympic tournaments, the original CFP format from 2014-2023, and NCAA Basketball. This analysis will first evaluate optimal Final 4 tournament designs, and expansions therefrom. In the traditional tournament design literature, the [1 versus 4, 2 versus 3] Final 4 format, as used in post-group play for certain Summer Olympic sports, the original CFP from 2014-2023, and the NCAA Women's and Men's Basketball Tournaments should the expected, or top-bracketed, Final Four emerge. This format is found to dominate all other single-elimination tournament designs in terms of efficacy and fairness [5]. As is often stated in economics, there is no easy *a priori* accounting for preferences (i.e., the attractiveness axiom). However, there is a clear accounting for the axioms of efficacy and fairness, and we will do so to compare classic and alternative tournament formats using discrete mathematical derivation, and optimization-based modeling techniques.

Namely, we pit the optimized, classic single-elimination format of [1 versus 4, 2 versus 3] against an increasingly high-profile hybrid format, the Page Playoff System (PPS), to determine if top-bracket double-elimination concessions can improve efficacy and fairness against the best-in-class single-elimination format. This remains an open, but increasingly important, question in the literature. In recent years, the PPS has been implemented in prominent sports leagues around the world, including the IPL, PSL, and BPL cricket playoffs and the NBA Play-In Tournament. The PPS may improve upon classic single-elimination formats by providing a top-bracket concession. We further consider the role of top-bracket concessions when used to expand a tournament, the CFP, from 4 to 12 teams. The CFP tournament design expanded to a non-binary number, necessitating that some teams—the top 4 teams in this case—receive a bye. Does this alternative top-bracket concession preserve or deteriorate efficacy and fairness?

By evaluating evolving tournament designs against their predecessors in this manner, we can help teams streamline efforts to create stable tournament designs that provide dynamic, meritorious incentives for teams to field the best teams possible.

## 2. Methods and Results

### 2.1 Four-team Knockout Playoffs

We first consider the 4-team single-elimination knockout playoff-design. Let  $A = [(1,4), (2,3)]$ ,  $B = [(1,2), (3,4)]$ , and  $C = [(1,3), (2,4)]$  be the distinct possible draws, where  $A$  is treated in the literature as the best classical, single-elimination draw with respect to the key tournament design criteria from the literature, as defined in the introduction.

The PPS (see Figure 1) is an alternative 4-team hybrid knockout playoff-design, used in professional cricket (Indian Premier League, Pakistan Super League, Bangladesh Premier League, NBA Play-In Tournament), curling, and softball. It provides an advantage to the top-two seeds in a Final 4 by providing them with a double-elimination safety net. The performance of the PPS relative to the optimized single-elimination format and with respect to key design criteria is heretofore unknown.

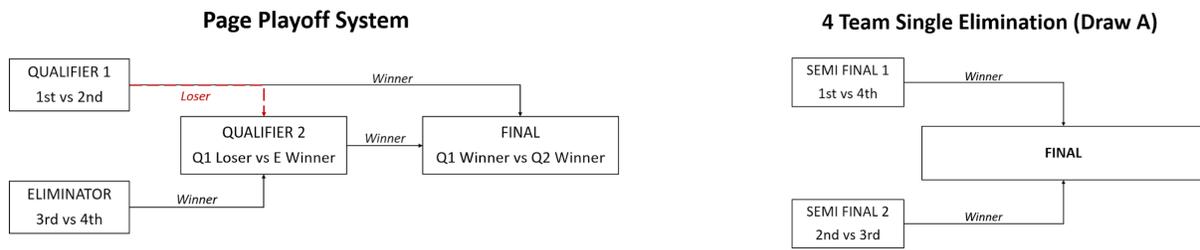


Figure 1: Page playoff system with double elimination for top bracket vs the famous 4 team single elimination design.

PPS is an elaboration of Draw B with a top-bracket double-elimination provision. PPS increases the chance of top-team(s) winning in four-team playoffs and therefore may overcome the design shortcomings of a standard Draw B format, as shown in the previous literature and extended herein.

### 2.1.1 Knockout Playoffs Analysis Methodology

Assume teams can be ranked best-to-worst and are numbered accordingly  $1, 2, \dots, n$ . Let  $P = (p_{ij})$  denote the probability-matrix, where  $p_{ij}$  represents the probability team  $i$  beats team  $j$ ,  $1 \leq i < j \leq n$ . The  $n \times n$  matrix  $P = (p_{ij})$  specifies the relative strength of the  $n$  teams. We make the following assumptions of  $P$ .

A1.  $0.5 \leq p_{ij} \leq 1$  for  $i \leq j$ , and  $0 \leq p_{ij} \leq 0.5$  for  $i \geq j$ . [The better team has a higher chance of winning a match between two teams.]

A2.  $p_{i,j} + p_{j,i} = 1$  [No Draws]

A3.  $p_{i,j}$  is non-decreasing in  $j$ , i.e.,  $P$  is non-decreasing along each row. [If team  $j$  is stronger than  $j'$ , then any third team  $i$  can more easily defeat  $j'$  than  $j$ ].

For a given probability matrix  $P$ , let  $q_j(X)$  denote the probability that team  $j$  wins under Draw  $X$ . For example,

$$q_1(A) = p_{14}(p_{23}p_{12} + p_{32}p_{13})$$

We consider fairness in the sense that the better a team is, the better chance it has of winning the tournament. Under this fairness definition, Draw  $X$  is fair if  $q_1(X) \geq q_2(X) \geq q_3(X) \geq q_4(X)$ .

**Theorem:** Assume A1, A2, and A3. Then among Draws A, B, C, and PPS:

(X1) PPS maximizes the probability of the best team winning the tournament, where A maximizes among classical designs A, B, and C.

(X2) A and PPS are fair (chance of winning tournament proportional to team strength).

[Proof is derived in the Appendix].

### 2.1.1 Knockout Playoff Design Simulation Results

Without loss of generality, let  $s_i$  be team  $i$ 's true strength, where  $s_1 \geq s_2 \geq \dots \geq s_n > 0$ . Using a Bradley-Terry statistical learning model, we define  $p_{ij} = s_i / (s_i + s_j)$ . This  $P$  satisfies the assumptions A1-3. Monte-Carlo simulations verify the Theorem's results by simulating 1 million tournaments. We considered  $s_i \sim U(a_i, b_i)$ , where  $(a_i, b_i)$  were chosen suitably to a) avoid overlap and have strictly different strengths (Figure 2) or b) to have some overlap to show robustness of the Theorem result (Figure 3). Figures 2 and 3 evaluate the likelihoods that teams win the tournament as a function of their team strength under PPS and Draw A, respectively. In Figure 2, we consider team strengths sampled from distinct, spaced-out uniform distributions, which means that the rank (seed) of the team has the same ordering as their team strength and the theorem conditions are satisfied.

#### Page Playoff System is Fair and Maximizes the Chance of Top Team Winning

Draw A is also fair but has smaller win probability, than Page playoff, for the top team.

A million tournament simulations with distinct spaced out strengths and theorem conditions satisfied.

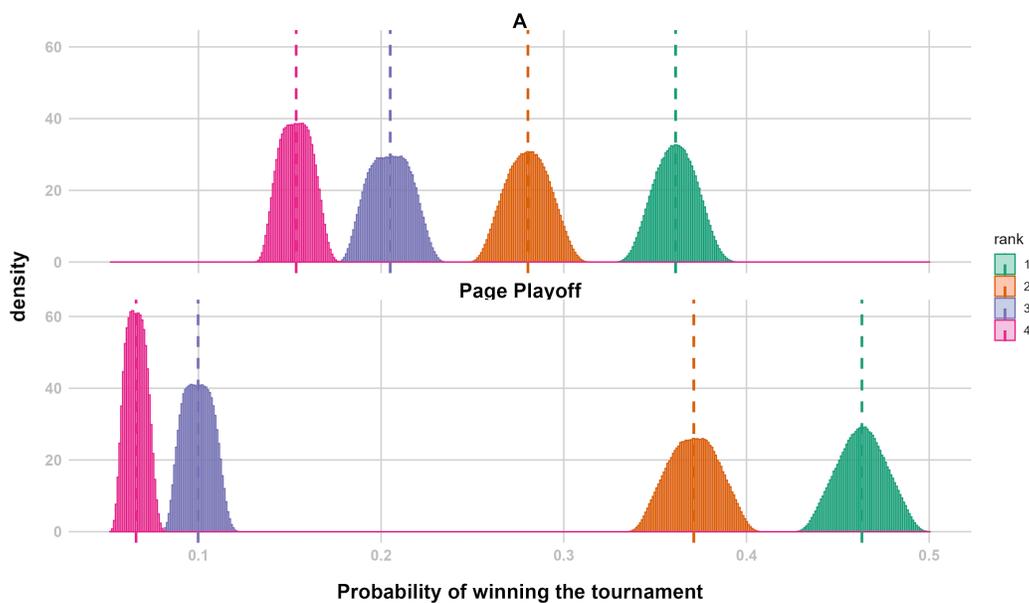


Figure 2: PPS and Draw A are fair and show high efficacy. Team strengths were sampled from distinct spaced out uniform distributions:  $U(a_1= 0.9, b_1 = 1)$ ,  $U(a_2= 0.85, b_2 = 0.75)$ ,  $U(a_3= 0.7, b_3 = 0.6)$ , and  $U(a_4= 0.575, b_4 = 0.5)$ .

Figure 2 shows that both PPS and Draw A are ordinally fair. That is, the probability of a team winning is directly ordinal to its rank or team strength. PPS shows much high efficacy than Draw A, however, by maximizing the chance of the top team(s) winning. Note that top teams are pulled sharply right, and bottom teams sharply left, for the same team strength draws from Draw A to PPS in Figure 2. Whereas every dog may have their day, it is difficult for bottom bracket dogs to repeatedly have their day when top bracket teams are provided double elimination. Therefore, top-bracket concessions raise efficacy even over the most efficacious single elimination Final 4 design.

Double elimination for the top 2 teams in PPS explains this shift. For example, a standard single-elimination Draw B is shown to be less efficacious than Draw A even. When Draw B is elaborated upon with top-bracket concessions, however, efficacy rises well above Draw A. In tournaments with longer initial rounds, such as round-robins, it is crucial to account for the performance of top-seeded teams, and PPS is a great choice for the playoff portion of such tournaments.

In the big-data simulation exercise of Figure 3, we let team strength source distributions overlap, which means the rank (seed) of the team isn't always the same as the order of teams based on strength (i.e., the possibility of inefficient seeding). Even under these inefficient seeding conditions, PPS demonstrates a markedly high level of efficacy relative to Draw A, with no distributional overlap between the probability of tournament victory for either of the top 2 teams by strength with either of the bottom 2 teams by strength. This shows that, under PPS, strong teams can retain a high chance of winning the tournament even if they are seeded (ranked) lower in the tournament bracket. That is, the PPS is compensatory of small-to-moderate seeding inefficiencies in a manner that Draw A, or any single-elimination draw, is not. For Draw A, it is interesting to note that, with overlapping team strengths, the win probability distributions shift closer to one another. This can potentially lead to relatively more upsets (lower ranked teams winning the tournament), which may relate to attractiveness for some fan bases. Additional data that links tournament upsets to viewership changes is needed to make this potential connection definitive.

#### Page Playoff System show Robustness

A million tournament simulations with overlapping team strengths. Note the overlap in win probability distributions.

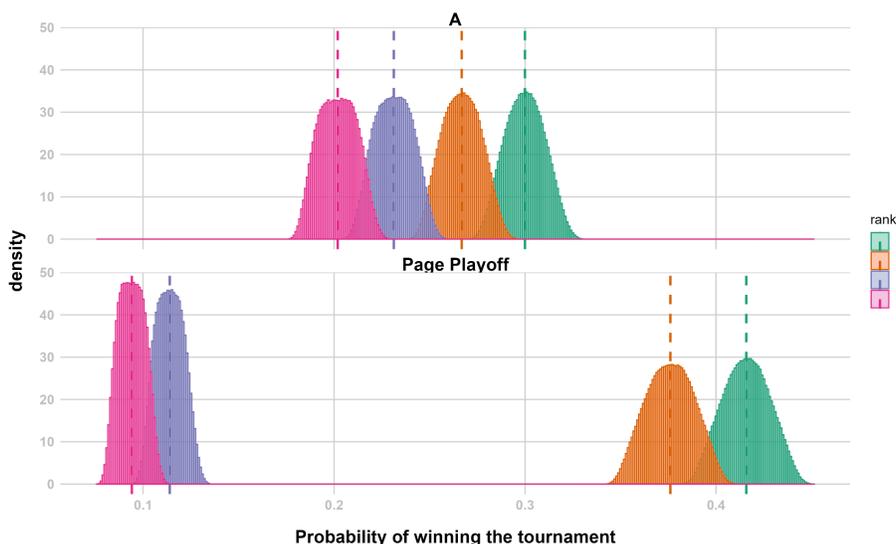


Figure 3: Overlapping team strengths show the robustness of PPS. Teams with higher strength but lower rank (seed) can still win. Uniform distributions:  $U(a_1= 0.9, b_1 = 1)$ ,  $U(a_2= 0.83, b_2 = 0.93)$ ,  $U(a_3= 0.75, b_3 = 0.85)$ , and  $U(a_4= 0.68, b_4 = 0.78)$ .

## 2.2 College Football Playoffs

The NCAA FBS football playoffs are among the main topics of discussion in U.S. collegiate athletics. The College Football Playoff (CFP) was established in the 2014 season to replace the Bowl Championship Series, aiming to provide a more definitive and equitable method for crowning a

national champion. Over its first decade (2014–2023), the format consisted of a four-team bracket selected by a committee. This design model often drew criticism for limiting tournament access to only a few dominant programs. Motivated by the need for broader inclusivity and increased engagement, the playoff expanded to a 12-team format beginning in the 2024 season. The expansion introduced on-campus, first-round games hosted by seeds 5–8, and granted first-round byes to the top-four seeds. The format has continued to evolve rapidly; while the 2024 model awarded byes to the top-four conference champions, the 2025 format is set to transition to straight seeding, where the top-four *ranked* teams receive byes regardless of conference championship status, further prioritizing team strength over conference affiliation. We will show that this was indeed the correct decision by the CFP.

Fan interest in the College Football Playoff (CFP) has remained high through these design iterations, with some ups-and-downs. Viewership peaked immediately with the inaugural 2015 Championship game between Ohio State and Oregon, which drew a record 34.1 million viewers. While ratings dipped to a low of 17.2 million for the uncompetitive 2023 Georgia-TCU blowout, they rebounded to 25 million for the 2024 final. The transition to the 12-team format in the 2024-25 season maintained strong engagement, with the Championship game drawing 22.1 million viewers despite scheduling challenges, while early-round, on-campus games generated significant local digital engagement lifts for participating schools. This devotion is perhaps best measured by the exorbitant costs fans are willing to absorb; average ticket prices for the National Championship have frequently topped \$2,000, with the 2025 Championship game shattering records at an average purchase price of approximately \$2,637. Beyond tickets, the expanded playoff format has intensified the travel burden, as fans making a deep postseason run now face cumulative travel costs that can exceed thousands of dollars for flights and lodging across multiple rounds.

### 2.2.1 12-Team Playoff Format

We analyze the new 12-team College football playoff format and consider changes from 2024-25-to-2025-26, particularly focusing on the conference champs getting a bye in the 2024 season. The 2024 season gave byes to the top 4 conference champions and a playoff spot for all power-5 conference champions, creating a scenario in which ASU, ranked 12, received a bye and the fourth seed. Also, Clemson, ranked 16, made the playoffs with the twelfth seed. The team rankings are announced by the CFP and, for our next simulation analysis of Figure 4, we assume they reflect the true strength of teams. Note that the CFP ranking may not always be based on true team strengths, and there are other objective methods to rank FBS teams, as we will show later.

Considering the nature of the sport and existing modeling literature [6], we assume a *Thurstone-Mosteller* stochastic-transitivity model,  $p_{ij} = \Phi(\theta_i - \theta_j + \eta)$ , where  $\theta_i$  is team  $i$ 's strength parameter,  $\eta$  accounts for home-advantage, and  $\Phi$  represents a standard normal CDF. We assume  $\theta = (\theta_1, \theta_2, \dots, \theta_n)$  follows a multivariate normal distribution. Strength parameter values  $\theta_i$  can be estimated using historical data and a Bayesian state-space model. However, we sampled these values from a multivariate normal distribution for this simulation. We simulate College Football Playoff iterations under 2025-26 rules, with and without home advantage, for Round 1 games, and under 2024-25 playoff-rules, where the four highest ranked conference champions received a first-round bye. To simulate the 2024-25 scenario, we randomly selected 4 teams from the 12 playoff teams.

### Heatmap of Win Probabilities

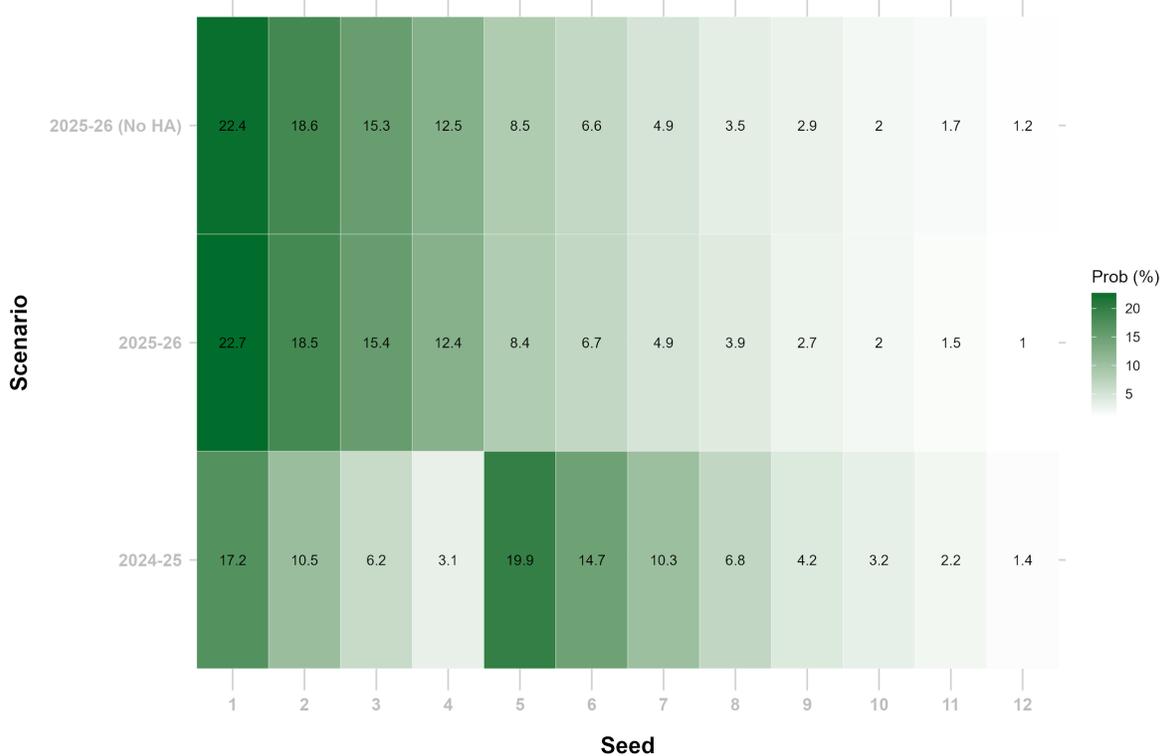


Figure 4: Note the degree to which the rule whereby conference champions received a bye and playoff spot affected the tournament win probabilities for the 2024 season simulation iterations. Some better teams, by strength, were pushed lower in seeding, and it shows. Results are from 50,000 simulated tournaments.

Figure 4 cell values are probability estimates (as percentages) of winning the tournament for given seed and tournament design, from 50,000 simulated tournaments. The tournament game outcomes are generated conditional on the simulated team strengths,  $\theta_i$ , with equally-spaced strengths and equal (moderate) uncertainty. As expected, the top 4 seeds have a significantly higher chance of winning the tournament, which makes winning one’s conference a high priority in the 12-team playoff. An interesting bi-modal pattern emerges for the 24-25 season, in which conference champions are seeded higher. Since conference champions could have been ranked lower but were moved to a higher seed, this seeding shift disrupts the probability distribution of the winner with respect to seed and lowers the efficacy and fairness of the tournament. The CFP made the correct decision to change the bye rule and award it to the top 4 ranked teams, regardless of conference champion status, in the 2025 season. This clearly and substantially raised efficacy and fairness according to the simulation results. With proactive tournament design optimization, the CFP may have avoided this inefficacious feature altogether.

### 2.2.2 College Football Rating System

Glickman and Stern [7] developed a Bayesian state-space model to estimate NFL team scores and game outcomes. The critical element of their approach is the estimation of team strengths, which are expected to vary over time. Their model accounts for this source of variability by modeling football game outcomes with a state-space model that assumes team-strength parameters follow a first-order autoregressive process. The model accounts for season-to-season changes driven by personnel shifts and other longer-term factors. Their model also incorporates a home-field advantage, while allowing for the possibility that its magnitude may vary across teams. The aim of the analysis is to obtain plausible inferences concerning team strengths and other model parameters, and to predict future game outcomes. The model details are presented in Figure 5.

The Glickman and Stern model approach provides a natural solution for estimating team strengths in College Football, as they rely on game-score differentials, account for home field advantage, and allow the team strength parameters to vary across years. Also, their framework provides a methodology to simulate game outcomes, which we will utilize later in the analysis.

### Modeling team ratings - Glickman & Stern (1998, 2017)

Let  $Y_{ijt}$  be the score differential in a game between teams  $i$  and  $j$  during year  $t$ :

$$Y_{ijt} \sim N( \overbrace{\mu_{ijt}}^{\text{game mean}}, \overbrace{\sigma^2}^{\text{game-level variance}} )$$

Simple model for game means with constant home-field advantage (HA):

$$\mu_{ijt} = \overbrace{\theta_{it} - \theta_{jt}}^{\text{difference in team ratings}} + \overbrace{\alpha}^{\text{constant HA}}$$

Autoregressive model for team ratings with constant innovation variance  $\tau^2$ :

$$\theta_{i(t+1)} \sim N(\gamma \cdot \theta_{i(t)}, \tau^2), \text{ where } 0 < \gamma < 1$$

$$\theta_{it_0} \sim N(0, \sigma_{team}^2)$$

Figure 5: Glickman and Stern model for team strength and game outcomes.

We use FBS regular season data from 2014 - 24 to model the Glickman team ratings (strength). Figure 6 shows the distribution of team rating (strength) grouped by conference over the years. SEC teams have dominated FBS for a decade. Figure 7 shows the top 12 teams by Glickman Rating for 2024. The Glickman ratings come with uncertainty estimates; ratings for different conferences with credible interval bands are presented in Figure 8.

## FBS Team Strength by Conference

Glickman Bayesian Ratings (Zero = Average FBS Team)

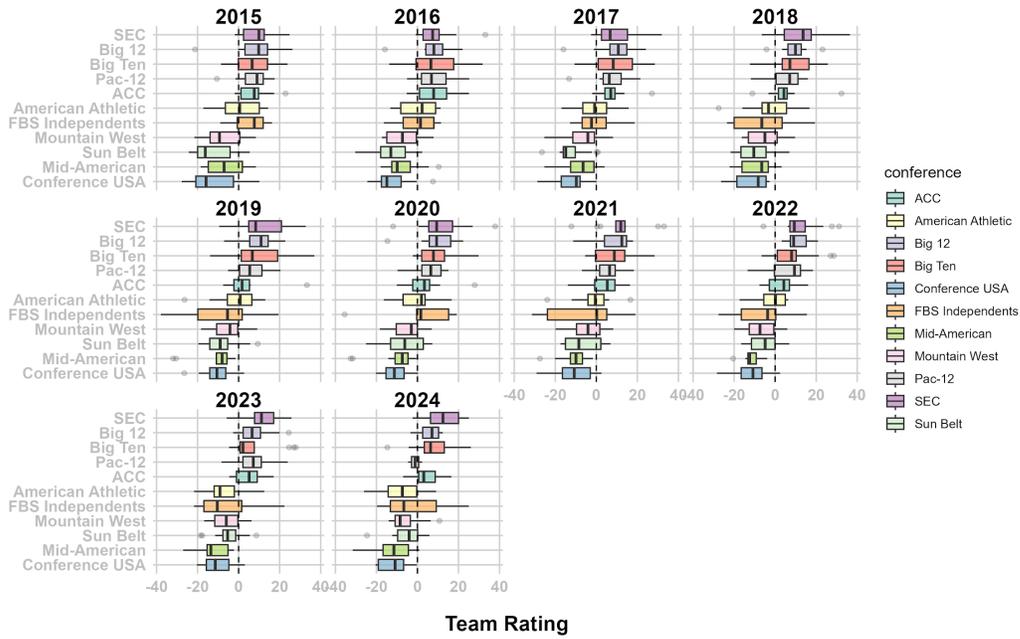


Figure 6: FBS team strengths grouped by conference over a decade. SEC, Big Ten and Big 12 teams have dominated the FBS since 2015.

## Rating Evolution: Top 12 Teams from 2024

9 teams were in playoffs. Ole Miss, South Carolina, and Miami weren't conference champs. Mean Glickman rating shown.

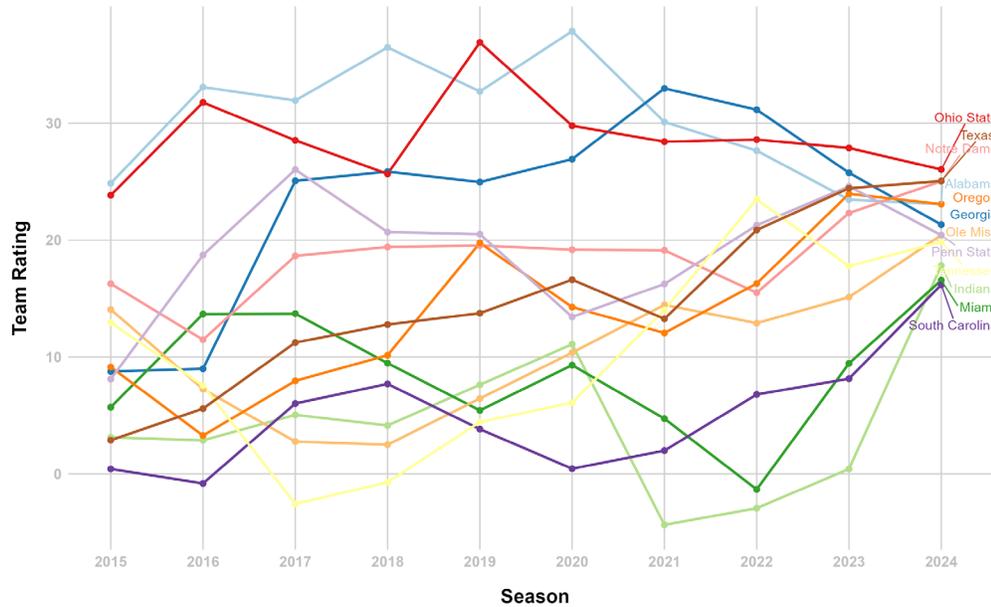


Figure 7: Top 12 teams based on Glickman rating, from 2024, are shown.

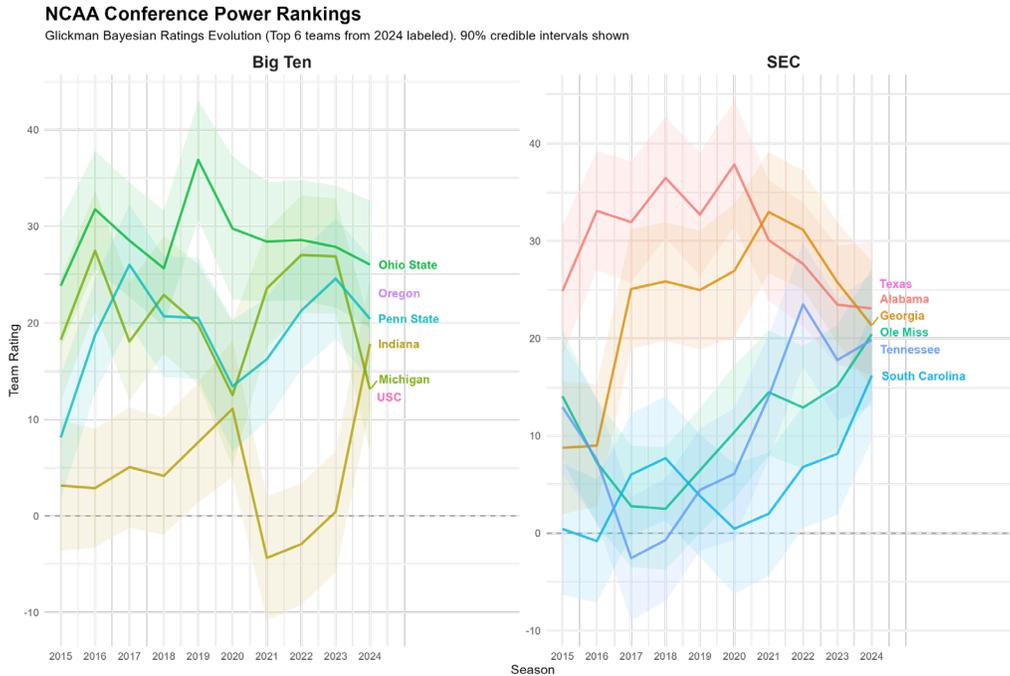


Figure 8: Glickman ratings for different conferences, evolution of ratings for the top 5 teams from 24 season.

Top teams, from the 2024 season, based on Glickman ratings are presented in Figure 7. Nine of those teams were in the playoffs, while the three teams that didn't make the playoffs were South Carolina, Mississippi and Miami. None of these three teams were Conference champions, which made it very difficult to reach the playoffs at that time. Note that Mississippi and South Carolina are from a very competitive SEC Conference.

In Figure 9, we compare the final 2024 CFP Top 25 rankings with the top 25 teams based on the Glickman Bayesian ratings for team strength. Note that the Glickman Bayesian model uses only score differential and home field advantage for the within conference regular season games. Also, these ratings are not adjusted to consider conference championship tickets to reach the playoffs. There is a strong harmony at the top of the CFP rankings, and 9 teams share the top 12 rank in both ranking systems. Clemson, ranked 16 in CFP and 15 in Glickman, reached the playoffs by virtue of a Conference Championship. Arizona State was also a Conference Champion.

## 2024 Playoff Consensus Check

Green = Team is inside the Top 12 in BOTH rankings

CFP RANK	TEAM	MODEL RANK	RATING	95% CI	MATCH_STATUS
1	Oregon	5	23.07	[16.64, 29.51]	Match
2	Georgia	6	21.32	[14.84, 27.88]	Match
3	Texas	2	25.06	[18.73, 31.37]	Match
4	Penn State	8	20.42	[14.06, 26.86]	Match
5	Notre Dame	3	25.04	[18.6, 31.31]	Match
6	Ohio State	1	26.04	[19.46, 32.68]	Match
7	Tennessee	9	19.84	[13.11, 26.61]	Match
8	Indiana	10	17.82	[11.13, 24.5]	Match
9	Boise State	-	-	-	Not in Model Top 25
10	SMU	13	15.70	[9.32, 22.03]	Mismatch
11	Alabama	4	23.08	[16.26, 29.88]	Match
12	Arizona State	25	11.01	[4.64, 17.49]	Mismatch
13	Miami	11	16.57	[9.68, 23.39]	Mismatch
14	Ole Miss	7	20.43	[13.72, 27.19]	Mismatch
15	South Carolina	12	16.18	[9.56, 22.82]	Mismatch
16	Clemson	15	14.50	[8.1, 21.01]	Mismatch
17	BYU	-	-	-	Not in Model Top 25
18	Iowa State	23	11.29	[4.45, 17.85]	Mismatch
19	Missouri	-	-	-	Not in Model Top 25
20	Illinois	-	-	-	Not in Model Top 25
21	Syracuse	-	-	-	Not in Model Top 25
22	Army	-	-	-	Not in Model Top 25
23	Colorado	21	11.78	[4.96, 18.64]	Mismatch
24	UNLV	-	-	-	Not in Model Top 25
25	Memphis	-	-	-	Not in Model Top 25

Figure 9: The table compares CFP and Glickman rankings, 2024 season. Interesting to note the harmony in selecting the top 12 teams.

## CFP History: Committee vs. Model

Green rows indicate teams that were Top 4 in BOTH rankings.

SEASON	TEAM	CFP RANK	MODEL RANK	FINAL STATUS
2023	Michigan	1	2	Champion
2023	Washington	2	14	Runner-Up
2023	Texas	3	5	Semifinalist
2023	Alabama	4	7	Semifinalist
2022	Georgia	1	1	Champion
2022	Michigan	2	4	Semifinalist
2022	TCU	3	11	Runner-Up
2022	Ohio State	4	2	Semifinalist
2021	Alabama	1	2	Runner-Up
2021	Michigan	2	4	Semifinalist
2021	Georgia	3	1	Champion
2021	Cincinnati	4	9	Semifinalist
2020	Alabama	1	1	Champion
2020	Clemson	2	3	Semifinalist
2020	Ohio State	3	2	Runner-Up
2020	Notre Dame	4	7	Semifinalist

2019	LSU	1	4	Champion
2019	Ohio State	2	1	Semifinalist
2019	Clemson	3	2	Runner-Up
2019	Oklahoma	4	6	Semifinalist
2018	Alabama	1	1	Runner-Up
2018	Clemson	2	2	Champion
2018	Notre Dame	3	8	Semifinalist
2018	Oklahoma	4	5	Semifinalist
2017	Clemson	1	3	Semifinalist
2017	Oklahoma	2	6	Semifinalist
2017	Georgia	3	5	Runner-Up
2017	Alabama	4	1	Champion
2016	Alabama	1	1	Runner-Up
2016	Clemson	2	5	Champion
2016	Ohio State	3	2	Semifinalist
2016	Washington	4	4	Semifinalist
2015	Clemson	1	4	Runner-Up
2015	Alabama	2	2	Champion
2015	Michigan State	3	15	Semifinalist
2015	Oklahoma	4	1	Semifinalist

Figure 10: We further back-test the CFP top 4 rankings against the Glickman model ratings. 64% of instances match up, where both committee and model pick the same teams for top 4.

Figure 10 shows the degree to which the model and CFP Committee matched up in selecting the top 4 (playoff) teams from 2015 - 23, 64% of teams matched. The conference championships could have impacted the Committee's decision in some cases, since in quite a few mis-matched cases the CFP and model rank are close. E.g. 2016 Clemson (CFP-2 vs Model-5), 2017 Georgia (CFP-3 vs Model-5), 2018 Oklahoma (CFP-4 vs Model-5), 2023 Texas (CFP-3 vs Model-5). To further validate the effectiveness of the Glickman model, we simulate the 2024 playoff using the model outputs. We supplied the bracket (teams) to the model, and the model used the team rating estimates to predict the outcome of each game along with the simulated scoreline and model confidence (probability) estimate. The model predicts all the outcomes correctly, meaning the actual outcome of games matched the predicted outcomes. Hence, the model correctly predicted the winner, OSU, with a simulated score of 26-25 (actual was 34-23) with model confidence of 53%. The Appendix contains the Glickman model's predictions for the 2025 playoff (Figure 12).

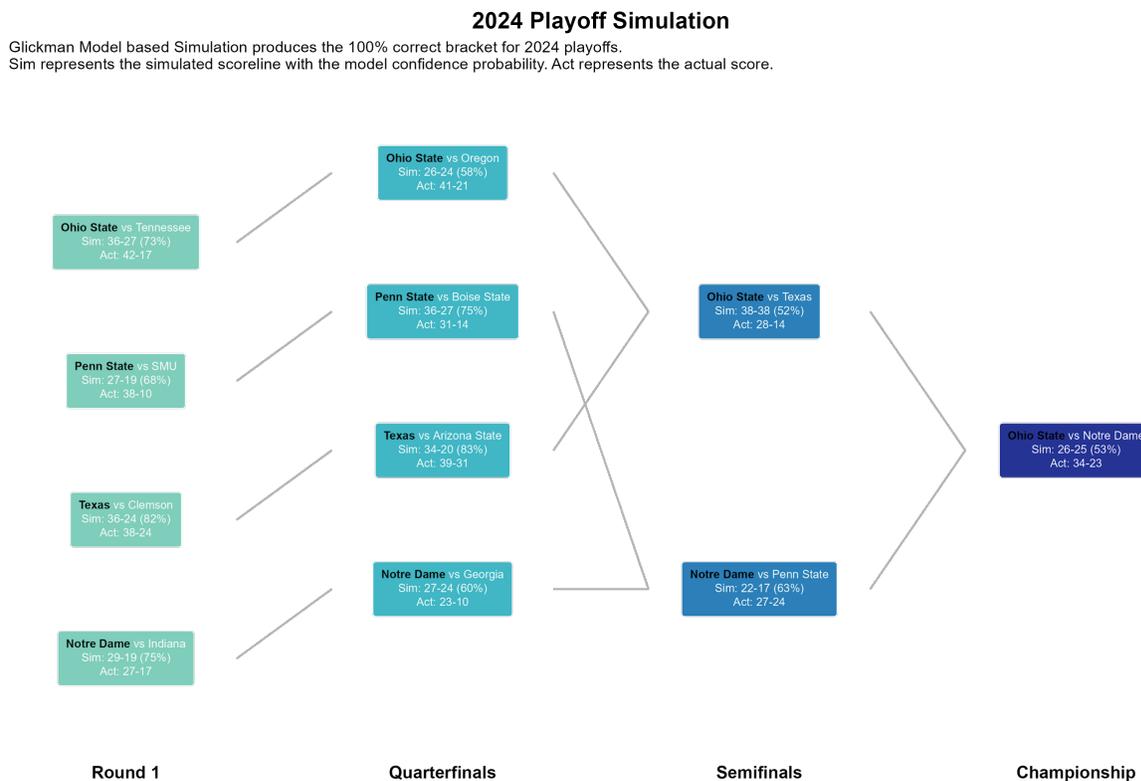


Figure 11: Glickman model predictions 100% match the actual outcomes of the 2024 playoffs, correctly predicting the winner. Sim represents the simulation predicted score with model confidence in % next to it. Act shows the actual score. See Appendix for 2025 playoff predictions, the brackets came out the day of submission deadline.

The Glickman model provides a comprehensive framework to estimate team performances in College Football. The model predictions align well with outcomes, making the model an excellent candidate for analysis support to CFP's decision making. The NCAA Soccer committee is known to use Ratings Percentage Index (RPI), along with other factors to inform their playoff selection

process. The NCAA Basketball Tournament Selection Committee uses Google's NET rating model. The CFP can effectively use an open source model, such as Glickman ratings, to assist their decision making and objectively decide edge cases.

### 3. Conclusion

This study demonstrates that tournament design can be optimized along impactful criteria or axioms using simulation and discrete mathematical analysis. In some cases, such optimization can explain unsatisfactory tournament designs that required years to correct, as in the case of the College Football Playoffs. Further, effective tournament design can provide teams with stronger incentives to field strong rosters toward playoff success. In turn, stronger play in a league can dynamically shift out demand in its own right. Indeed, tournament and competition format design are chief analytic functions of leagues.

We also find in our study that other recently mainstream hybrid tournament designs eclipse the optimized, classic, single-elimination designs in terms of efficacy and fairness. A comparison of alternative designs illuminates the trade-offs between fairness and efficacy. We prove a theorem showing the fairness and high efficacy properties of the Page Playoff System. When contrasting the Page Playoff System with the standard Draw A bracket, distinct structural incentives emerge. The Page Playoff System minimizes the risk of premature elimination for top seeds by offering a "double chance" mechanism, thereby strongly rewarding consistent, season-long excellence. The specification of a Page Playoff System by the NBA (Play-In Tournament) and leading professional cricket leagues improves the meritorious nature of playoff outcomes in those leagues. The Page Playoff System is particularly well-suited to tournaments with longer initial rounds (such as round robins). The traditional Draw A is fair and is still a viable alternative when extra games are not feasible in the playoffs.

Our analysis demonstrates that the College Football Playoff's 2025 transition to "straight seeding" corrects a critical inefficiency in the 2024 CFP format. The 2024 model's prioritization of Conference Champions created a bimodal distortion that effectively penalized high-performing non-champions. By removing these constraints, the 2025 format restores a meritocratic distribution of success probabilities. Additionally, our analysis confirms that the vast advantage of winning one's Conference is ascending to the playoffs and receiving a first-round bye. Furthermore, we demonstrate the effectiveness of modeling frameworks, such as the Glickman model, for analyzing team performance and appropriately rating teams for playoff consideration. In the future, the CFP would be well-served to follow NCAA Basketball and Soccer Selection Committees by seeking structured statistical model support. This would allow the CFP to minimize the design pitfalls and growing pains it has experienced throughout its existence.

Ultimately, as postseason formats grow in complexity and financial scale, structural decisions should no longer rely solely on intuition or tradition. This research underscores the need to integrate rigorous probabilistic modeling into the design process. By conducting appropriate simulation analyses that vary constraints, administrators can empirically balance the competing demands of entertainment value, revenue, and competitive equity, ensuring that the tournament's architecture honors the integrity of the sport and the demands of fans.

## References

- [1] Solntsev, Ilya & Kurov, Andrey. (2024). Sports Tournament Design as a Source of Economic Impact (Through the Example of Russian Premier League). *Journal of Applied Economic Research*, 23. 304-340. [10.15826/vestnik.2024.23.2.013](https://doi.org/10.15826/vestnik.2024.23.2.013).
- [2] Reilly, B., & Witt, R. (2021). The Effect of League Design on Spectator Attendance: A Regression Discontinuity Design Approach. *Journal of Sports Economics*, 22(5), 514-545. <https://doi.org/10.1177/1527002521989393>
- [3] Devriesere, K., Csato, L., & Goossens, D.R. (2024). Tournament design: A review from an operational research perspective. *Eur. J. Oper. Res.*, 324, 1-21.
- [4] Szymanski, S. (2003). The economic design of sporting contests. *Journal of Economic Literature*, 41(4):1137-1187.
- [5] Horen, J., & Riezman, R. (1985). Comparing draws for single elimination tournaments. *Operations Research*, 33(2), 249-262.
- [6] Hennessy, J.P., & Glickman, M.E. (2015). Bayesian optimal design of fixed knockout tournament brackets. *Journal of Quantitative Analysis in Sports*, 12, 1 - 15.
- [7] Glickman, M. E., & Stern, H. S. (1998). A State-Space Model for National Football League Scores. *Journal of the American Statistical Association*, 93(441), 25-35. <https://doi.org/10.1080/01621459.1998.10474084>

## Appendix

Check next page for 2025 playoff predictions and theorem proof.

## 2025 Playoff Simulation Results

Glickman Model based simulation results for 2025 playoff. FBS regular season data from 2015 - 2025.  
 Predicted on 7th Dec, 2025

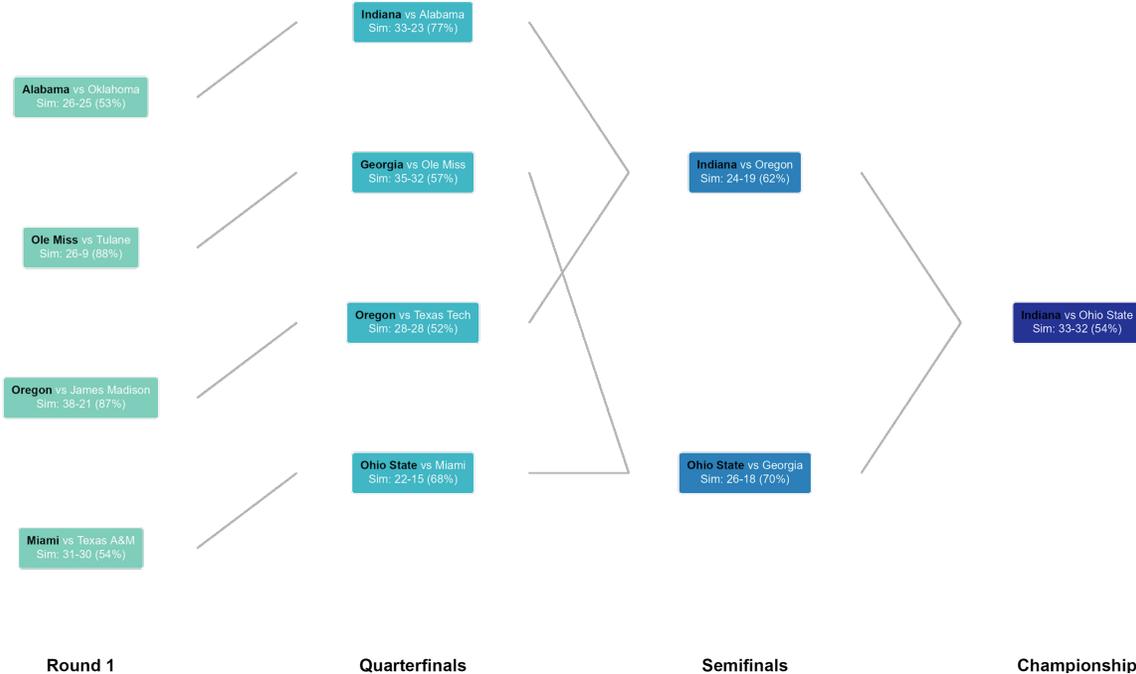


Figure 12: The Glickman model's 2025 playoff prediction. Indiana winning against Ohio State, with a prediction for a close game. The simulation was conducted on Dec 7th, the day CFP announces the playoff bracket and the paper submission deadline.

Theorem Proof: Draw A maximizes the probability of best team winning among draws A, B, and C. See [1]. To show PPS maximizes the probability of best team winning, we just need to show that  $q_1(\text{PPS}) \geq q_1(A)$ , which is same as  $q_1(\text{PPS}) - q_1(A) \geq 0$ .

$q_1(\text{PPS}) - q_1(A) = (p_{12}p_{34}p_{23}p_{12} + p_{12}p_{34}p_{32}p_{13} + p_{12}p_{43}p_{24}p_{12} + p_{12}p_{43}p_{42}p_{14} + p_{21}p_{34}p_{13}p_{12} + p_{21}p_{43}p_{14}p_{12}) - (p_{12}p_{14}p_{23} + p_{13}p_{14}p_{32})$ . Below we will show that this difference is non-negative.

The proof for fairness is excluded for brevity. But the arguments are the same as for the best team winning (efficacy).

Write the left big sum (all capital  $P$ 's) as

$$S := \underbrace{P_{12}^2 P_{34} P_{23} + P_{12} P_{34} P_{32} P_{13} + P_{21} P_{34} P_{13} P_{12}}_{\text{(terms multiplied by } P_{34})} + \underbrace{P_{12}^2 P_{43} P_{24} + P_{12} P_{43} P_{42} P_{14} + P_{21} P_{43} P_{14} P_{12}}_{\text{(terms multiplied by } P_{43}=1-P_{34})}$$

So the quantity we want to prove nonnegative is

$$\Delta := S - (P_{12} P_{14} P_{23} + P_{13} P_{14} P_{32}).$$

Step 1 — replace the "reverse" probabilities by complements.

Use  $P_{ji} = 1 - P_{ij}$  (assumption (2)) where needed; in particular

$$P_{21} = 1 - P_{12}, \quad P_{32} = 1 - P_{23}, \quad P_{42} = 1 - P_{24}, \quad P_{43} = 1 - P_{34}.$$

Step 2 — group terms by the factor  $P_{34}$  and  $P_{43} = 1 - P_{34}$ . Define

$$A := P_{12}^2 P_{23} + P_{12} P_{32} P_{13} + P_{21} P_{13} P_{12}, \quad B := P_{12}^2 P_{24} + P_{12} P_{42} P_{14} + P_{21} P_{14} P_{12}.$$

Then

$$S = P_{34} A + P_{43} B.$$

So

$$\Delta = P_{34} A + P_{43} B - (P_{12} P_{14} P_{23} + P_{13} P_{14} P_{32}).$$

Step 3 — extract a common factor  $P_{12}$  from  $A$  and  $B$ . Both brackets contain a factor  $P_{12}$ , so write  $A = P_{12} A_1$  and  $B = P_{12} B_1$  where

$$A_1 = P_{12} P_{23} + P_{32} P_{13} + P_{21} P_{13}, \quad B_1 = P_{12} P_{24} + P_{42} P_{14} + P_{21} P_{14}.$$

Thus

$$\Delta = P_{12} (P_{34} A_1 + P_{43} B_1) - (P_{12} P_{14} P_{23} + P_{13} P_{14} P_{32}).$$

Step 4 — compare term-by-term using the monotonicity ordering (assumption (3)).

Assumption (3) gives, for fixed row index and increasing column index, nondecreasing probabilities. In particular (with the natural ordering  $1 < 2 < 3 < 4$ ) we have

$$P_{12} \leq P_{13} \leq P_{14}, \quad P_{23} \leq P_{24} \leq P_{34}.$$

Also from (1) and (2) we know every  $P_{ij} \in [0, 1]$  and the complements are nonnegative.

Now inspect the two bracketed expressions  $A_1$  and  $B_1$ . Because  $P_{13} \geq P_{12}$  and  $P_{24} \geq P_{23}$ , each appearance of a "later" column index increases the corresponding factor:

– In  $A_1$ , the terms are  $P_{12} P_{23}$ ,  $P_{32} P_{13}$ ,  $P_{21} P_{13}$ . Replacing  $P_{21} = 1 - P_{12}$  and  $P_{32} = 1 - P_{23}$  shows each summand is nonnegative, and the presence of the factor  $P_{13}$  (which is  $\geq P_{12}$ ) in two summands makes  $A_1$  at least as large as the single product  $P_{14} P_{23}$  once it is multiplied by the appropriate prefactor  $P_{34}$  or  $P_{12}$  (because  $P_{14} \geq P_{13} \geq P_{12}$  and  $P_{34} \geq P_{23}$ ). Concretely, monotonicity implies that replacing  $P_{13}$  by  $P_{14}$  (or  $P_{12}$  by  $P_{13}$ ) does not decrease such products.

– Similarly for  $B_1$ : its three summands contain  $P_{24}$  or  $P_{14}$  or  $P_{12}$  in combinations and, by monotonicity, are at least as large as the combinations that appear in the subtracted pair  $P_{12} P_{14} P_{23} + P_{13} P_{14} P_{32}$  once correctly weighted by  $P_{34}$  and  $P_{43}$ .

Putting these monotonicity comparisons together (each term in  $P_{34} A_1 + P_{43} B_1$  is a nonnegative combination of factors that are coordinatewise  $\geq$  the corresponding factors in  $P_{12} P_{14} P_{23} + P_{13} P_{14} P_{32}$ ), we conclude

$$P_{34} A_1 + P_{43} B_1 \geq P_{14} P_{23} + P_{13} P_{32}.$$

Multiplying both sides by the nonnegative scalar  $P_{12}$  gives

$$P_{12} (P_{34} A_1 + P_{43} B_1) \geq P_{12} P_{14} P_{23} + P_{12} P_{13} P_{32}.$$

Finally observe  $P_{12} P_{13} P_{32} \geq P_{13} P_{14} P_{32}$  is not generally true, but the earlier more detailed termwise comparison takes care of those cross-terms — the bracketed combination  $P_{34} A_1 + P_{43} B_1$  was designed to dominate exactly the two-subterm combination that is subtracted. Therefore

$$\Delta \geq 0,$$

as required.