# Dynamic Opponent Choice in Tournaments

**Nicholas G. Hall**

**September 15, 2021**

**schedulingseminar.com**

THE OHIO STATE UNIVERSITY FISHER COLLEGE OF BUSINESS

# Tournament Design

- There is a long history of research on tournament design in both the operations research and economics literatures.
- But, perhaps surprisingly, it is not an easy or "well solved" problem.
- In fact, some *impossibility* results show that several reasonable performance criteria cannot be achieved simultaneously or even individually.
- Even beyond that issue, there are various *randomly occurring deficiencies* in tournament design.

# Objectives of Tournaments

- Provide the players with equally fair opportunities (measurable in several ways)
- Select among the players, rank them, and reward them according to their performance
- Provide an appealing event for spectators, to make the tournament financially successful
- Motivate the players to perform well at all times.

# Tournament Type

- Our focus is on tournaments that include:

- A preliminary stage (sometimes called a "group stage") that determines a *ranking* or *seeding* for a subset of the players who continue to later rounds.

- Followed by two or more rounds of single elimination play, organized using information from this ranking.

Examples include:

- Playoffs in U.S. major professional sports (NFL, MLB, NBA, NHL, MLS)

- NCAA men's and women's basketball championships

- ATP men's and women's tennis tournaments

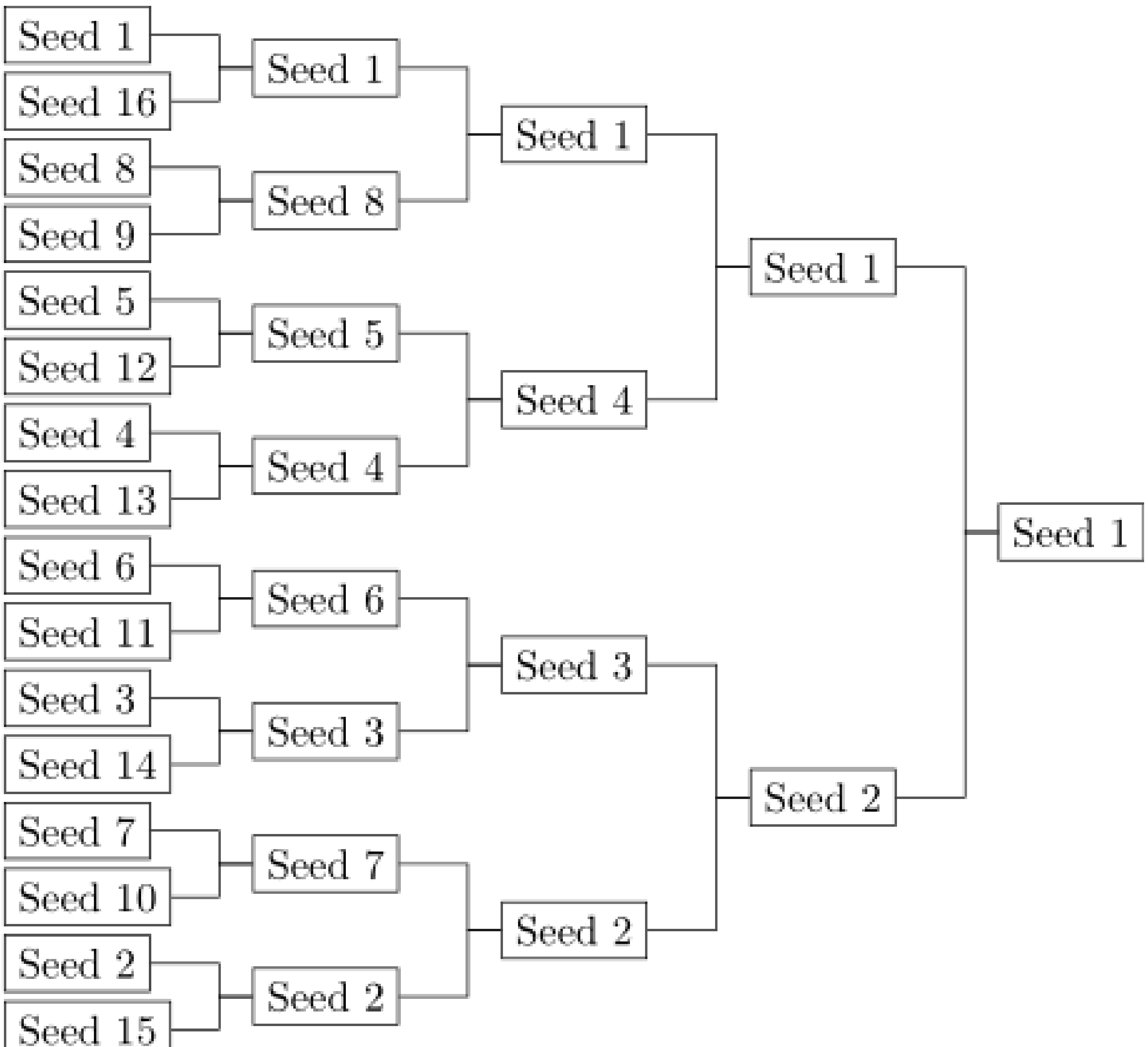
- FIFA World Football Cup, ICC World Cricket Cup

- Various tournaments outside of sports (bridge, chess, …).

# Definitions

- Since our work applies to both individual and team tournaments, we use *player* to describe either an individual player or a team.
- The *preliminary* or *group stage* may extend over an entire season to establish a ranking.
- Alternatively, a ranking may be based on a *statistical measure* or a *committee decision*.
- The definition of single elimination is that defeat immediately eliminates a player from the tournament, but there can be *multiple games* (e.g., the NBA playoffs and the MLB World Series are both single elimination).

# Conventional Single Elimination Tournament Design




This is a *binary tree* that is *symmetric* (without *byes*), or a *bracket*. But many other brackets are possible. For example, swap games [8,9] and [2,15]. Each bracket or "seeding" has different measures of fairness, for example tournament win probabilities for different players.

# Impossibility Results for Tournament Design Performance (1 of 3)

- *Fairness Under Medium Ranking Condition:* If the players can be ranked such that each player ranks all the lower ranked players in the same sequence according to probability of beating them, then Player ***i*** (who is expected to beat Player ***j***) should have a tournament win probability at least as high as Player ***j*** does.

- Intuitively, each row of the pairwise win probability matrix is nondecreasing.

- For a tournament with *eight* players, no bracket achieves this fairness measure for an arbitrary pairwise win probability matrix (Horen and Riezman 1985). [Proof by finding a counterexample matrix for all 315 available brackets.]

# Impossibility Results for Tournament Design Performance (2 of 3)

- *General Fairness Under Strong Ranking Condition:* If for each pair of Players ***i*** and ***j>i***, ***i*** has a <u>higher probability</u> of beating every other player than ***j*** does, then ***i*** should have a tournament win probability at least as high as ***j*** does.
- <u>Intuitively</u>, each row of the pairwise win probability matrix is nondecreasing *and* each column is nonincreasing ("Player ***i*** dominates Player ***j***.")
- For a tournament with *eight or more players*, <u>no bracket</u> achieves this fairness measure for an arbitrary pairwise win probability matrix (Vu and Shoham 2011). [Theoretical proof.]

# Impossibility Results for Tournament Design Performance (3 of 3)

- *Elimination of Shirking:* Players should have a nonnegative incentive to win every game. For a tournament with fixed groupings at the following round, allowing exactly one player to continue from each group is both necessary and sufficient to eliminate *shirking* (or, "*tanking*"), i.e. deliberately failing to win (Vong 2017). [Note that this solution would frustrate spectators and reduce interest in the tournament.]

# Deficiencies in Tournament Design

- Top ranked players may randomly incur "bad matchups" against other players, which introduces an unnecessary element of luck.

- Being ranked very highly after the preliminary round does not provide any particular advantage.

- The use of a conventional fixed bracket fails to allow players to take into account information that develops during the tournament, such as injuries to other players.

- The design encourages shirking at the preliminary round, in order to achieve an easier path through the tournament.

We propose a new tournament design to address these issues.

# Shirking

- Diminishes the prestige, credibility and profitability of the tournament (and perhaps the players), and in some cases national pride.
- Due to substantial amounts of gambling, for example $1.8b on the 2018 FIFA World Football Cup, shirking raises the possibility of legal liability.
- Since shirking is more likely to be used by top players, disqualifying them is a poor solution.
- Prevents the tournament from providing an unbiased ranking of the players for public interest and for use at future events.

# Famous Examples of Shirking

- The 2005-06 L.A. Clippers lost late season games to avoid the Dallas Mavericks, resulting in changes to NBA playoff design.
- In the 2006 Winter Olympics ice hockey competition, the Swedish coach publicly discussed losing against Slovakia, to avoid playing the Czech Republic or Canada later. Sweden lost 3-0 and won the gold medal.
- At the 2012 Summer Olympics, the Chinese, Indonesian and S. Korean women's badminton teams were all disqualified for deliberately losing their group stage matches.
- At the 2018 FIFA World Football Cup, the winner of the last group match between England and Belgium could face Brazil in the quarterfinal. Both teams rested their top players, and Belgium played their players out of position. Belgium won 1-0 but later beat Brazil.

# England Score Big Win By Losing To Belgium

Billy Haisley
6/28/18 5:21PM • Filed to: ENGLAND



Photo: Matthias Hangst (Getty)

In today's crucial group stage match between England and Belgium, the Brits narrowly managed to come out on top by coming out on bottom of a hard-fought 1-0 loss. With this phenomenal loss, England won entry to the much weaker side of the knockout round bracket, and now have a clear lane to the

# Even Worse Shirking

- In the 1998 Tiger Cup Asian football competition, the winner of the last group game between Thailand and Indonesia would need to travel to Hanoi to play the hosts Vietnam in front of a hostile crowd, whereas the loser could conveniently stay in Ho Chi Minh City and play Singapore, which was perceived as a weaker team.

- Here is what happened (Indonesia in white, Thailand in red) …

# Even Worse Shirking

# Even Worse Shirking

- In the 1998 Tiger Cup Asian football competition, the winner of the last group game between Thailand and Indonesia would need to travel to Hanoi to play the hosts Vietnam in front of a hostile crowd, whereas the loser could conveniently stay in Ho Chi Minh City and play Singapore, which was perceived as a weaker team.

- The Indonesian player Mursyid Effendi, who egregiously scored for Thailand at the end of the video, was banned from international soccer for life.

# Literature

Probability that Player *i* beats Player *j*

Horen and Riezman (1985) compare different brackets for single elimination tournaments with four and eight players.

They assume that the $N$ players' pairwise win probability matrix satisfies:
(i) $0.5 \leq p_{ij} \leq 1$ for $1 \leq i < j \leq N$,
(ii) $0 \leq p_{ij} \leq 0.5$ for $1 \leq j < i \leq N$,
(iii) $p_{ij} + p_{ji} = 1$ for $1 \leq i, j \leq N$, and
(iv) $p_{ij}$ is nondecreasing in $j$ for $1 \leq i, j \leq N$.

A pairwise win probability matrix that satisifies conditions (i) – (iv) is *strongly stochastically transitive* (David 1963), or SST.

This is described above as a *medium ranking* condition.

**Remark:** if every pairwise win probability matrix was SST, then the tournament design problem would be easier (*but still not very easy*) to solve!

# Literature

Horen and Riezman (1985) establish several *reasonableness criteria* for tournaments with $N$ players. They assume that the win probability matrix is SST.

(i) does the design maximize the probability that the best player wins?

(ii) does it maximize the probability that the best two teams meet?

(iii) is it order preserving, i.e., no stronger player has a lower probability of winning than a weaker player?

If $N = 4$, there are three distinct brackets, and the conventional bracket matchup is the unique one that satisfies these criteria.

If $N = 8$, there are 315 distinct brackets. Criterion (i) is satisfied by eight of them. However, for criterion (iii), no bracket is satisfactory.

This negative result motivates our study. We will apply the above three performance criteria.

# SST or Not?

Djokovic has beaten Nadal in 29 out of 55 matches.

| | R. Nadal | | R. Federer | | A. Murray | |
|---|---|---|---|---|---|---|
| N. Djokovic | [29–26] | **.527** | [27–23] | **.540** | [25–11] | **.694** |
| R. Nadal | | | [24–16] | **.600** | [17–7] | **.708** |
| R. Federer | | | | | [14–11] | **.560** |

Win Probability Matrix: Men's Professional Tennis Rivalries, 4 Players.

The above matrix is SST. But if we add a fifth player …

| | R. Nadal | | R. Federer | | A. Murray | | J.M. Del Potro | |
|---|---|---|---|---|---|---|---|---|
| N. Djokovic | [29–26] | **.527** | [27–23] | **.540** | [25–11] | **.694** | [16–4] | **.800** |
| R. Nadal | | | [24–16] | **.600** | [17–7] | **.708** | [11–6] | **.647** |
| R. Federer | | | | | [14–11] | **.560** | [18–7] | **.720** |
| A. Murray | | | | | | | [7–3] | **.700** |

Win Probability Matrix: Men's Professional Tennis Rivalries, 5 Players.

The last matrix is not SST. This is because Djokovic has a better record against Del Potro than against Murray, but Nadal has the reverse.

This example provides insight about "bad matchups" – Del Potro is a bad matchup for Nadal (relative to their overall levels of play).

# More Extreme Bad Matchup

| | $P2$ | $P3$ | $P4$ |
|---|---|---|---|
| $P1$ | 1 | 1 | 0 |
| $P2$ | | 0 | 0 |
| $P3$ | | | 1 |

Extreme Bad Matchup for Player P1.

A conventional bracket would match $P1$ and $P4$, to the disadvantage of player $P1$.

The only matchup from which $P1$ can win the tournament is if it plays $P2$, which enables $P3$ to eliminate $P4$.

This seems unfair to the highest ranked player, $P1$.

We have a proposal to address this problem ...

# We Propose an Alternative Design to Solve this Problem

Let the players, in ranked order based on their previous performance, **choose their opponents**

- from those *still available*, and
- assuming they were not themselves *previously chosen*,

at each round!

# How New Is This Design (1 of 3)?

Existing applications:

- Austrian Ice Hockey League (EBEL)
- Southern Professional Hockey League (U.S.).
- Canadian and U.S. Bridge Federations
- Chess PRO League
- Sailing
- Some e-games.

# How New Is This Design (2 of 3)?

ns Cities

InsideHook

## MLB Considering Controversial Reality TV-Inspired Playoff Expansion

**MLB's new plan would allow higher-seeded wild-card teams to choose opponents**

The MLB logo on batting practice balls. (Jill Weisleder/MLB Photos via Getty)

BY EVAN BLEIER / FEBRUARY 11, 2020 9:00 AM

**Sports News, 2020-02-11**

# How New Is This Design (3 of 3)?

Existing literature:

- Apparently no *published* research work.

- November 2019 working paper by Julien Guyon (Courant Institute, NYU): looks at the *tournament's objective* of maximizing the number of home games for players, provides an application to a 2020 European soccer competition, does not use any pairwise win probability information or prove any anti-shirking results.

# Questions to Ask about the Proposed Opponent Choice Design

- Does it resolve the random “bad matchup” problem?

- Does it provide reasonable results for real world sports data?

- Does it improve on the bracket design for the three reasonableness criteria, under large sample testing and sensitivity analysis?

- Does it reduce shirking?

# Comments about this Design

1. Allowing opponent choice opens up a <u>wide range of new designs and related analysis</u> in tournament design, and our work is an early exploration of this range.

2. We will <u>compare the tournament results</u> under our design, using the three reasonableness criteria discussed above, with those under a conventional bracket.

3. The tournament design changes we are recommending occur at the single elimination stage of the tournament; however, they reduce the incentive for strategic behavior such as shirking at the preliminary stage, thereby <u>potentially improving both stages</u>.

# Assumptions

The players have a ranking, or total order, based on earlier performance or in some cases the decisions of a seeding committee.

Each player has either *full* or *partial knowledge* of the pairwise win probability matrix.

At every round, each player which can choose its opponent does so with the objective of *maximizing its tournament win probability*.

# Full Information Assumption

First, we will study a situation where each player has complete and accurate information about the pairwise win probability matrix.

This is most realistic where the players have played each other frequently, as for example in professional tennis.

# Resolving Ties for Ranking

Our tournament design requires that the players enter the single elimination stage of the tournament with an empirically determined ranking, without ties.

This ranking is *weak*, for example the matrix is in general not SST.

Major sports have detailed tie-breaking rules in place: the NFL uses 12 sequential performance-based rules, and the NBA uses a similar procedure.

Tournaments that use limited within-group performance information, including the FIFA World Football Cup, could have more ties for ranking.

Tie-breaking options in that case include the use of a pre-tournament ranking or more detailed group performance information.

# Addressing the Bad Matchup Problem

Let $q_i, i = 1, \ldots, N$, denote the probability that Player $i$ wins the tournament.

| | $P2$ | $P3$ | $P4$ |
|---|---|---|---|
| $P1$ | $p_{12}$ | $p_{13}$ | $0$ |
| $P2$ | | $p_{23}$ | $p_{24}$ |
| $P3$ | | | $p_{34}$ |

Example of One Difficult Opponent.

If $P1$ chooses $P2$, then $q_1 = p_{12}p_{34}p_{13}$.
If $P1$ chooses $P3$, then $q_1 = p_{13}p_{24}p_{12}$.
Therefore, $P1$ chooses $P_2$ iff $p_{12}p_{34}p_{13} \geq p_{13}p_{24}p_{12}$, or $p_{34} \geq p_{24}$.
So, $P1$ lets the other player *which has the better chance* to beat $P4$ do so.

**Remark:** Player 1's choice of player 2 or 3 as opponent is *independent* of whether $p_{12}$ or $p_{13}$ is larger. Hence, a myopic strategy is not generally optimal, even for $N = 4$.

# How Many Players Will Choose?

At the semifinal round, only one player chooses its opponent, since this defines both the matches to be played.

At the quarterfinal round, only three players choose their opponents, for the same reason.

At the round-of-16, only seven players choose their opponents, for the same reason.

In general, for a tournament round with $N$ players, the number of players who choose their opponents is $N/2 - 1$.

**Remark:** Our tournament design allows flexibility for
(a) a smaller number of players to choose their opponents, or
(b) higher ranked players to be exempt from being chosen.

# Our Tournament Design

Consider a generic tournament with $n$ rounds and $N = 2^n$, $n = 2, 3, \ldots,$ players.

At any round, Player 1 freely chooses its opponent.

Then, the highest ranked unchosen player, i.e., 2 if not chosen by 1, or 3 otherwise, makes its choice, and so on.

In total, in the first round $N/2 - 1$ players choose their opponents sequentially, an $(N/2 - 1)$-stage Stackelberg game (Fudenberg and Tirole 1993).

# Static vs. Dynamic Ranking

In our tournament design, the *ranking* of the players is important, since it determines the order of opponent choice at each round.

Under a *static* ranking, a player's ranking remains the same as when it entered the single elimination stage of the tournament.

Under a *dynamic* ranking, the winner of a game inherits the higher ranking of the two players in that game for the next round.

**Example:** Suppose $P8$ beats $P1$ and $P7$ beats $P2$ at the same round. Then, $P8$ is ranked first, and chooses its opponent first, at the next round.

More generally, various ranking rules can be adopted, for example interpolating between the two models we consider, for example the Elo system (Wikipedia 2019) used to determine chess ratings.

# Algorithm Opponent Choice

*Input*

$p_{ij}$, for $i, j = 1, \ldots, N$, $N = 2^n$, $i \neq j$. Let $\mathcal{S}_a$ denote the set of all the players.

*Value Function*

For a given set $\mathcal{S}$ of players with relative rankings $1, \ldots, |\mathcal{S}|$ and respective original rankings $(1), \ldots, (|\mathcal{S}|)$, let $Q(\mathcal{S}) \equiv \{q_{\mathcal{S},(1)}, \ldots, q_{\mathcal{S},(|\mathcal{S}|)}\}$ denote the corresponding tournament win probabilities of the players, given all players' optimal choices of opponents.

*Boundary Condition*

When $|\mathcal{S}| = 2$, i.e, only two players with relative rankings 1 and 2, we have $q_{\mathcal{S},(1)} = p_{(1)(2)}$ and $q_{\mathcal{S},(2)} = p_{(2)(1)}$, where (1) and (2) are the two players' original rankings. Hence we assume that, for any $\mathcal{S} \in \mathcal{S}_a$, $|\mathcal{S}| = 2, 4, \ldots, N/2$, we have found $Q(\mathcal{S}) = \{q_{\mathcal{S},(1)}, \ldots, q_{\mathcal{S},(|\mathcal{S}|)}\}$.

**Remark:** The algorithm starts by enumerating all possible final rounds, then all possible semifinal rounds, and so on. Within each round, subsets of players are evaluated from smallest to largest.

# Algorithm Opponent Choice

*Recurrence Relation*

Let set $\mathcal{U}$ contain all the players with undecided opponents, where player $j$ has the highest ranking in $\mathcal{U}$ and $\mathbf{X} \cup \mathbf{X}' \cup \mathcal{U} = \mathcal{S}_a$. Then, player $j$ chooses its opponent to maximize its tournament win probability $q_j = q_{\mathcal{S}_a,j}$, where:

Probability of subset $S$

$$q_{\mathcal{S}_a,j} = \max_{i \in \mathcal{U}} \sum_{\mathcal{S} \in \Omega(\mathbf{X} \cup \{j\}, \mathbf{X}' \cup \{i\})} p_{\mathcal{S}}\, q_{\mathcal{S},j}$$

Probability of win, given subset $S$

where $\Omega(\mathbf{X} \cup \{j\}, \mathbf{X}' \cup \{i\})$ is known by previous calculation.

**Theorem:** Algorithm Opponent Choice finds the optimal sequence of opponent choices for all $N = 2^n$ players, and their tournament win probabilities, in $O(2^{\sum_{h=2}^{n} 2^{h-1}} \prod_{l=2}^{n} \prod_{h=1}^{2^{l-1}} (2^l - 2h + 1))$

**Remark:** For 16 players, as for single elimination at the FIFA World Football Cup, the algorithm solves the problem in 12 hours on a laptop computer.

# Real World Data

Big Data!

Sports Hub Data (2019): ATP men's tournaments from 1991–2016.
Includes 95,360 tennis matches over 1,902 tournaments.
Focus on 28,534 matches potential matches between top-16 seeds (round-of-16 or later). There are 5,943 such matches, won by the higher (respectively, lower) seeded player in 3,788 (resp., 2,155) instances.
Results of these 5,943 matches estimate win probabilities between all pairs of seeded players, based on win frequency.
**Example:** No.1 seeds have a 97-78 record against No.2 seeds, for an estimated win probability for the No.1 seed of $97/175 = .554$.
The win probability matrix is shown in the following table.

| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | .554 | .657 | .665 | .727 | .713 | .763 | .721 | .763 | .795 | .783 | .813 | .806 | .886 | .813 | .773 |
| 2 | | .669 | .690 | .621 | .630 | .603 | .670 | .784 | .650 | .650 | .844 | .718 | .833 | .817 | .773 |
| 3 | | | .436 | .549 | .617 | .626 | .578 | .650 | .438 | .633 | .548 | .703 | .583 | .735 | .690 |
| 4 | | | | .503 | .506 | .667 | .636 | .533 | .750 | .412 | .538 | .733 | .593 | .607 | .815 |
| 5 | | | | | .484 | .607 | .333 | .512 | .667 | .600 | .634 | .688 | .500 | .667 | .545 |
| 6 | | | | | | .519 | .560 | .448 | .567 | .596 | .643 | .500 | .560 | .643 | .722 |
| 7 | | | | | | | .409 | .583 | .617 | .579 | .565 | .385 | .545 | .500 | .583 |
| 8 | | | | | | | | .483 | .532 | .643 | .379 | .385 | .455 | .583 | .706 |
| 9 | | | | | | | | | .583 | .455 | .615 | .500 | .545 | .583 | .385 |
| 10 | | | | | | | | | | .545 | .706 | .357 | .545 | .545 | .583 |
| 11 | | | | | | | | | | | .500 | .583 | .583 | .385 | .357 |
| 12 | | | | | | | | | | | | .615 | .667 | .667 | .357 |
| 13 | | | | | | | | | | | | | .500 | .545 | .417 |
| 14 | | | | | | | | | | | | | | .500 | .500 |
| 15 | | | | | | | | | | | | | | | .583 |

Win Probability Matrix for Seeded Tennis Players.

# Example of the Algorithm (1 of 2)

We calculate seed No.1's tournament win probability if it chooses to play against seed No.7, which is its optimal choice.
Then, the other quarterfinal games, given optimal sequential selections by the next highest ranked available players which are seeds Nos.2 and 3, respectively, are $(2,4),(3,6)$ and $(5,8)$.
The probability that the winners of the four quarterfinal games are seeds Nos.1, 2, 3 and 5 is $p_{17} \times p_{24} \times p_{36} \times p_{58} = .763 \times .690 \times .617 \times .333 = \mathbf{.108}$.
On the condition that the other semifinal round players are seeds Nos.2, 3 and 5, seed No.1 has three choices:
Play seed No.2: $q_1 = .554[(.549 \times .657) + (.451 \times .727)] = 0.382$.
Play seed No.3: $q_1 = .657[(.621 \times .554) + (.379 \times .727)] = 0.407$.
Play seed No.5: $q_1 = .727[(.669 \times .554) + (.331 \times .657)] = \mathbf{0.427}$. ← Best choice

Since Algorithm Opponent Choice operates by backward recursion, these probabilities are known to seed No.1 at the time of choosing its quarterfinal opponent.

In view of seed No.1's objective of maximizing its tournament win probability, it chooses to play seed No.5 at the semifinal round.

# Example of the Algorithm (2 of 2)

The probabilities of the eight possible semifinal configurations in which seed No.1 is included appear in the third column of the table as $p(s_i)$.

The *conditional probability* for seed No.1 to win the tournament from each semi-final appears in the fourth column as $p(q_1|s_i)$.

The fifth column shows the *joint probability* $p(q_1, s_i) = p(s_i) * p(q_1|s_i)$ of the semifinal configuration $i$ and seed No.1 winning the tournament.

Finally, the probability of seed No.1 winning the tournament appears at the bottom of the fifth column, where $q_1 = \sum_{i=1}^{8} p(q_1, s_i)$.

| $i$ | $s_i$ | $p(s_i)$ | $p(q_1\|s_i)$ | $p(q_1, s_i)$ |
|---|---|---|---|---|
| 1 | $(1, 5, 2, 3)$ | **.108** | **.427** | .046 |
| 2 | $(1, 8, 2, 3)$ | .217 | .424 | .092 |
| 3 | $(1, 5, 2, 6)$ | .067 | .445 | .030 |
| 4 | $(1, 8, 2, 6)$ | .134 | .442 | .059 |
| 5 | $(1, 5, 4, 3)$ | .049 | .481 | .023 |
| 6 | $(1, 8, 4, 3)$ | .097 | .477 | .046 |
| 7 | $(1, 5, 4, 6)$ | .030 | .500 | .015 |
| 8 | $(1, 8, 4, 6)$ | .060 | .496 | .030 |
| | | | **Total** | **.342** |

Tournament Win Probability of Seed No.1 if it Chooses Seed No.7.

# Results for Top Tennis Seeds

| Seed $i$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Opponent $i$ | 4 | 3 | 2 | 1 |
| $q_i$ | .391 | .353 | .124 | .132 |

Tournament Win Probabilities for Seeded Tennis Players at Semifinal Round.

Our dynamic and static designs and the bracket design all give the same results when N = 4.

| Seed $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Opponent | 7 | 4 | 6 | 2 | 8 | 3 | 1 | 5 |
| $q_i^D$ | .342 | .242 | .101 | .078 | .049 | .064 | .039 | .085 |
| $q_i^S$ | .342 | .242 | .107 | .078 | .050 | .064 | .037 | .080 |
| $q_i^B$ | .315 | .201 | .121 | .109 | .081 | .072 | .044 | .056 |

Tournament Win Probabilities for Seeded Tennis Players at Quarterfinal Round.

Our dynamic and static designs both increase the tournament win probabilities of seeds Nos.1 and 2 when N = 8, which is a desirable feature and can be viewed as a reward for a high ranking.

# Results for Top Tennis Seeds

| Seed $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Opponent | 14 | 12 | 15 | 16 | 13 | 11 | 10 | 9 |
| $q_i^D$ | .337 | .223 | .086 | .080 | .056 | .047 | .031 | .036 |
| $q_i^S$ | .336 | .232 | .089 | .085 | .057 | .047 | .030 | .035 |
| $q_i^B$ | .313 | .228 | .073 | .079 | .056 | .050 | .035 | .037 |
| | | | | | | | | |
| Seed $i$ | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| Opponent | 8 | 7 | 6 | 2 | 5 | 1 | 3 | 4 |
| $q_i^D$ | .023 | .016 | .020 | .009 | .012 | .009 | .009 | .007 |
| $q_i^S$ | .021 | .014 | .020 | .008 | .010 | .005 | .006 | .006 |
| $q_i^B$ | .025 | .017 | .021 | .013 | .013 | .011 | .010 | .018 |

Tournament Win Probabilities for Seeded Tennis Players at Round-of-16.

The main winner from our tournament design is the first-mover seed No.1. Seeds Nos.3 and 4 benefit from our design under the dynamic ranking, and seeds Nos.2, 3, 4 and 5 benefit from our design under the static ranking.

The tournament win probability of Seed No.14 reduces from .011 under the bracket design to .005 under the static ranking, but only to 0.009 under the dynamic ranking.

# Computational Study

- We study the results of our tournament design using the three performance criteria established by Horen and Riezman (1985):

- The probability that the top ranked player wins the tournament

- The probability that the top two players meet

- Order preservation: the tournament win probabilities of the players preserve their original ranking.

# Sensitivity to Matrix Irregularity

We begin with a highly regular pairwise win probability matrix with $N = 2^n$ players where $p_{ij} = .5 + .05(j - i)$, for $1 \leq i < j \leq N$.

| | Player 2 | Player 3 | Player 4 |
|---|---|---|---|
| Player 1 | .55 | .60 | .65 |
| Player 2 | | .55 | .60 |
| Player 3 | | | .55 |

Highly Regular Win Probability Matrix for $N = 4$.

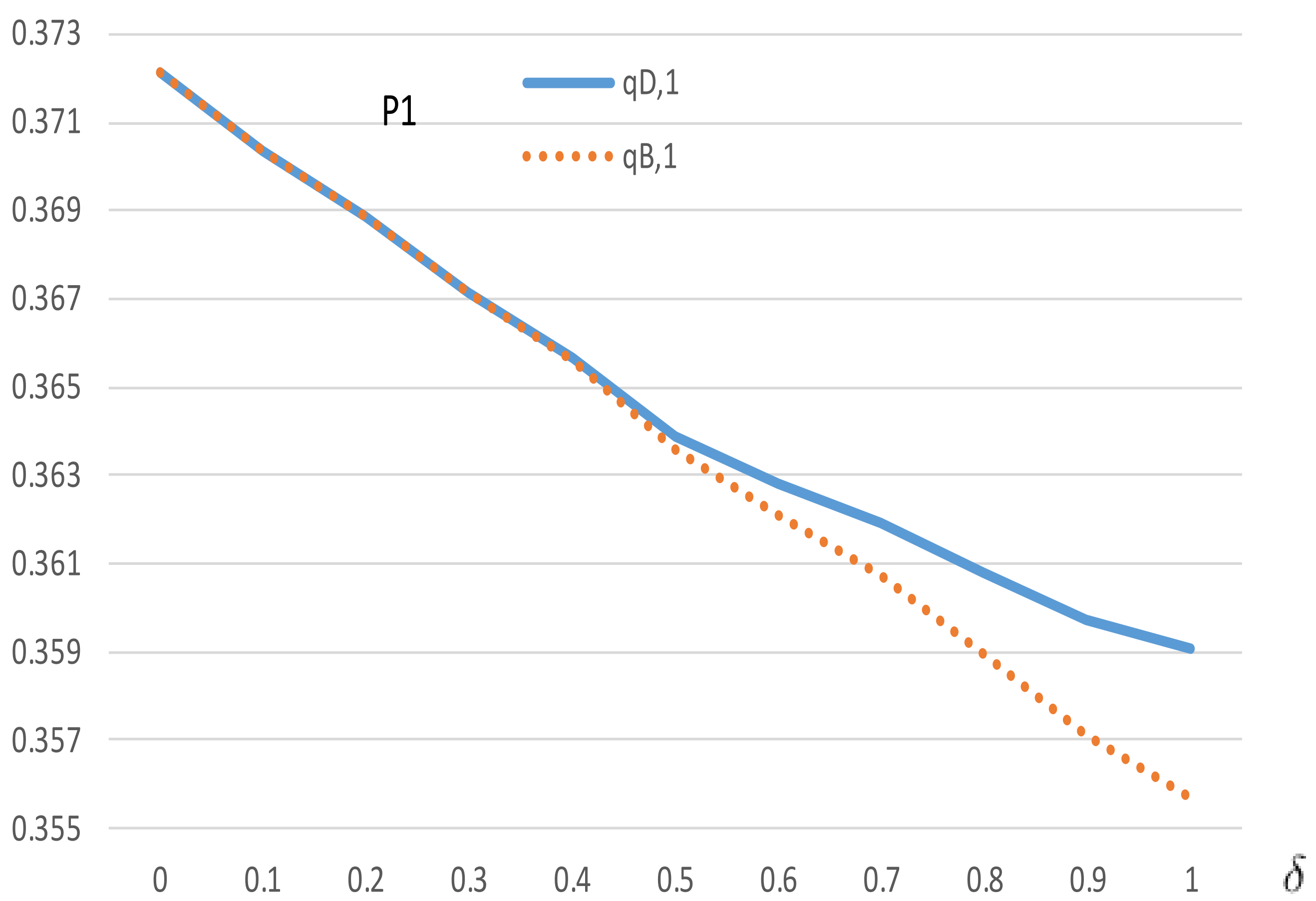
Observe that this matrix has a *strong ranking*, where Player $i$ dominates Player $j$, for $1 \leq i < j \leq N$.

# Sensitivity to Matrix Irregularity

We also introduce an amount of irregularity $\delta$ by adding to each matrix entry a randomly generated value $d \sim U[-\delta, +\delta]$, for $\delta = 0, .01, .02, \ldots, .1$.

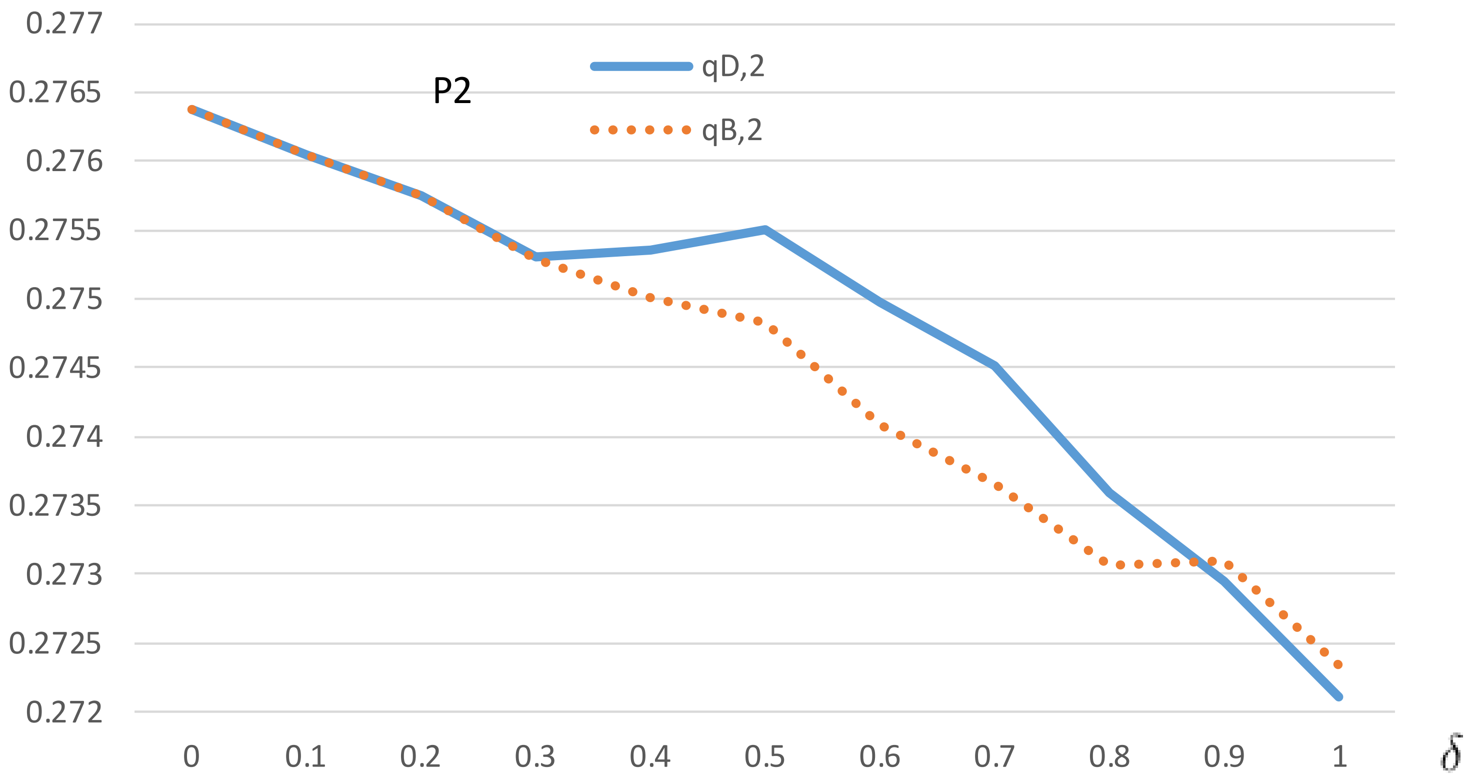
We generate the value of $d$ independently for each matrix entry. For the semi-final and quarterfinal rounds with $\delta > 0$, we have 6 and 28 random variables, respectively, and hence use a large sample size of 10,000 for each $\delta$.

For the semifinal round, the dynamic and static rankings are the same in our tournament design since only Player 1 can choose its opponent, whereas for the quarterfinal round they are different.



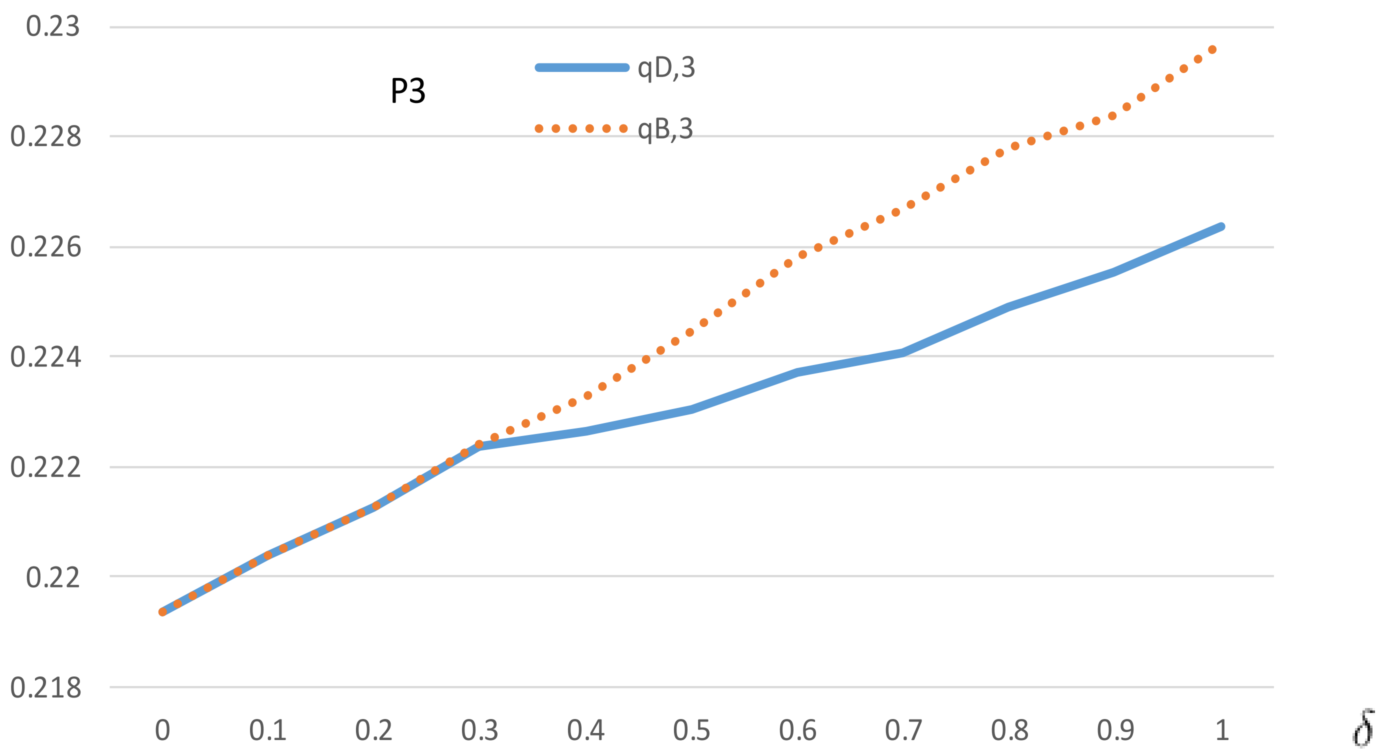

**Semifinal Round: Player 1**

Decreasing tournament win probability due to less dominance as matrix irregularity increases. The largest benefit of our design occurs at highest irregularity.



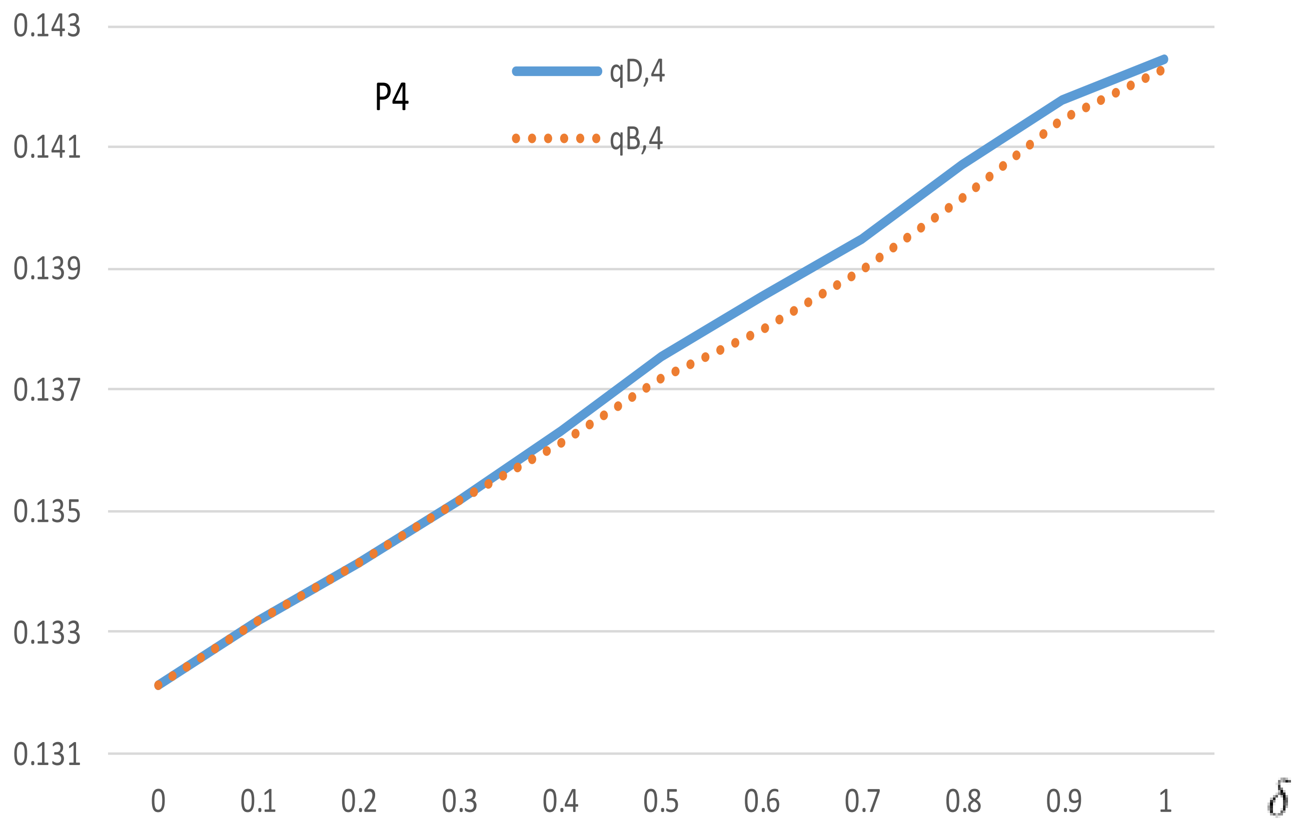

## Semifinal Round: Player 2

Decreasing tournament win probability due to less dominance as matrix irregularity increases. The largest benefit of our design occurs at moderate irregularity, when Player 1 chooses Player 3.




**Semifinal Round: Player 3**

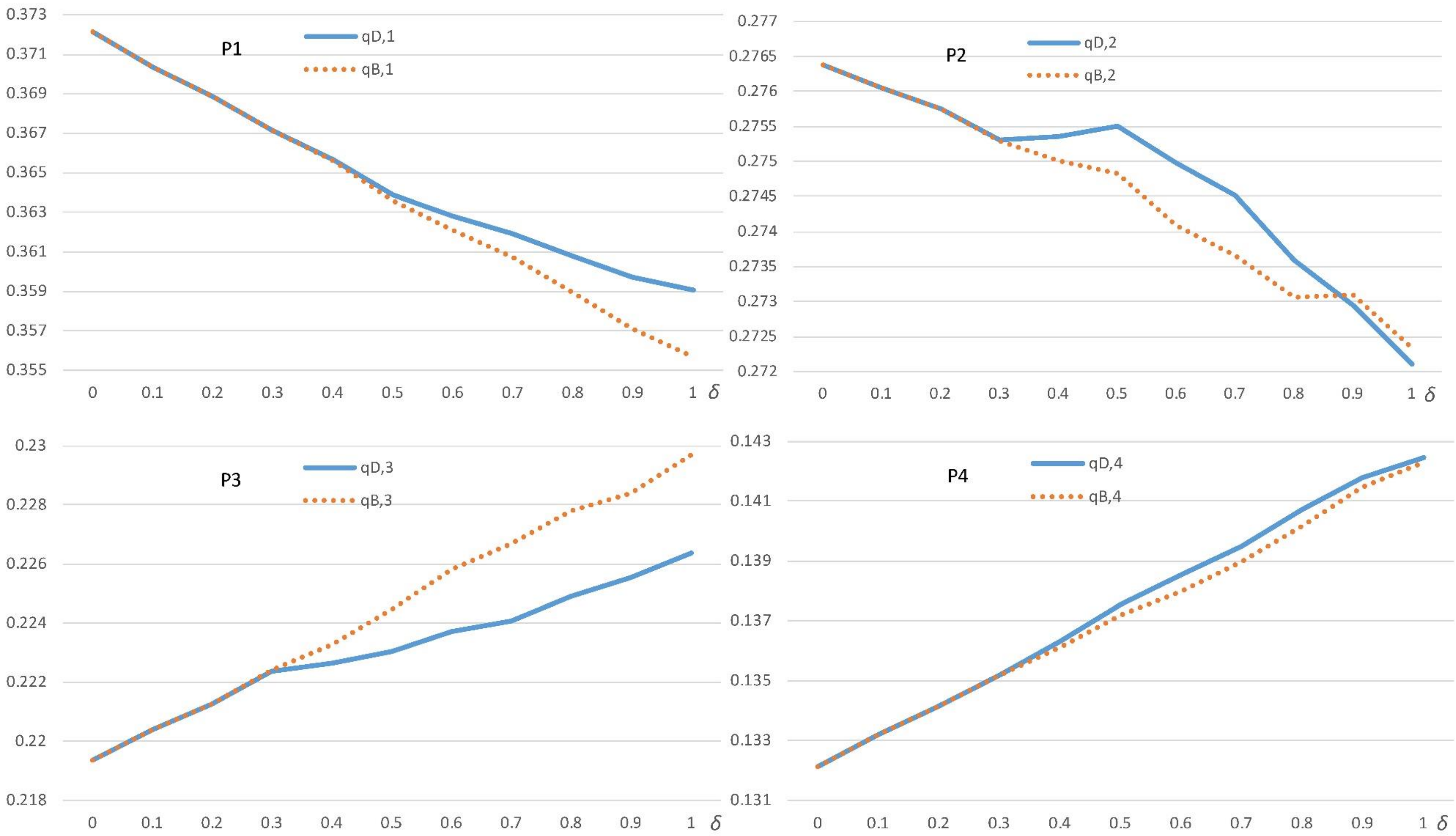
Increasing tournament win probability due to less dominance by the top players as matrix irregularity increases. Loses from our design due to increased probability of being chosen by Player 1.

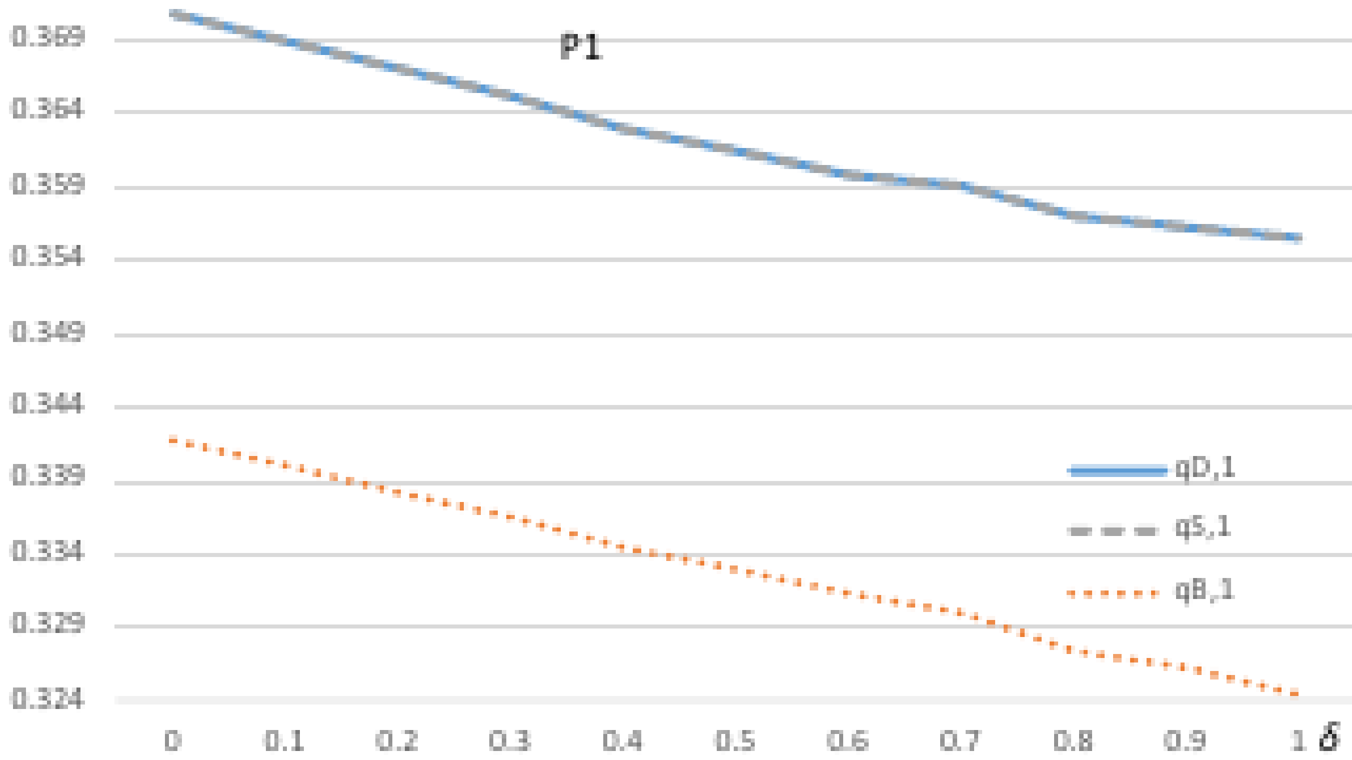



**Semifinal Round: Player 4**

Increasing tournament win probability due to less dominance by the top players as matrix irregularity increases. Gains slightly from our design, due to decreased probability of being chosen by Player 1.

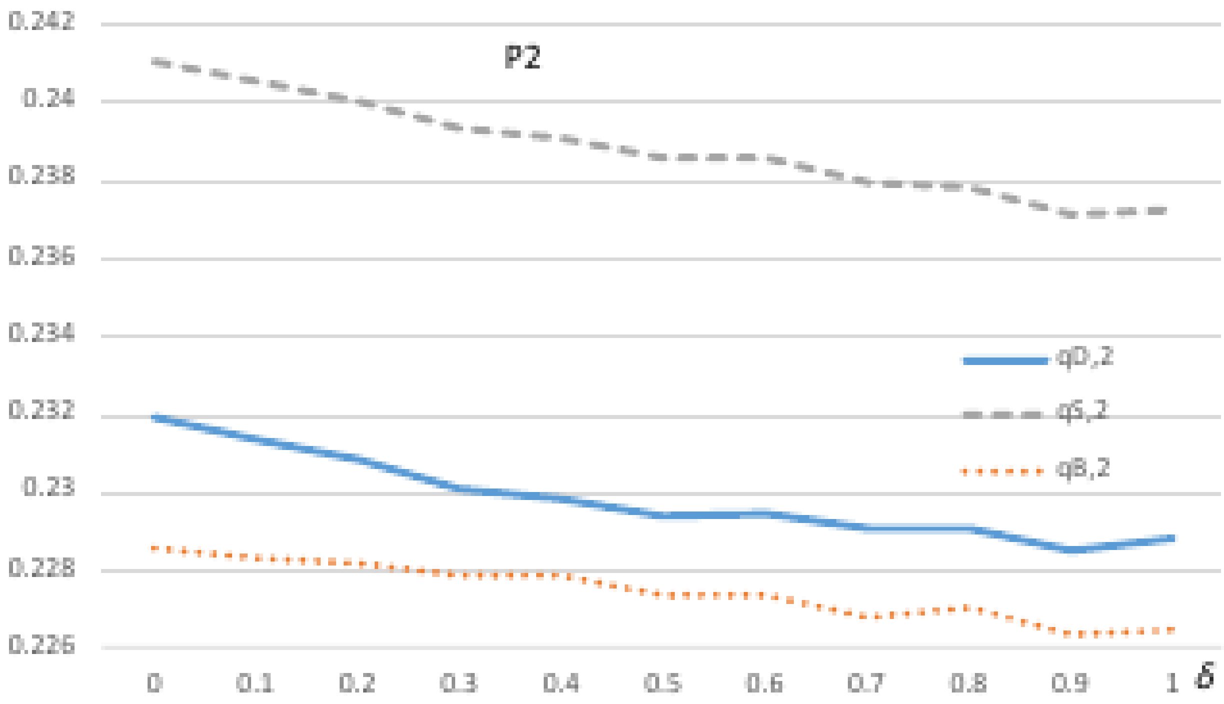
# Semifinal Round Results





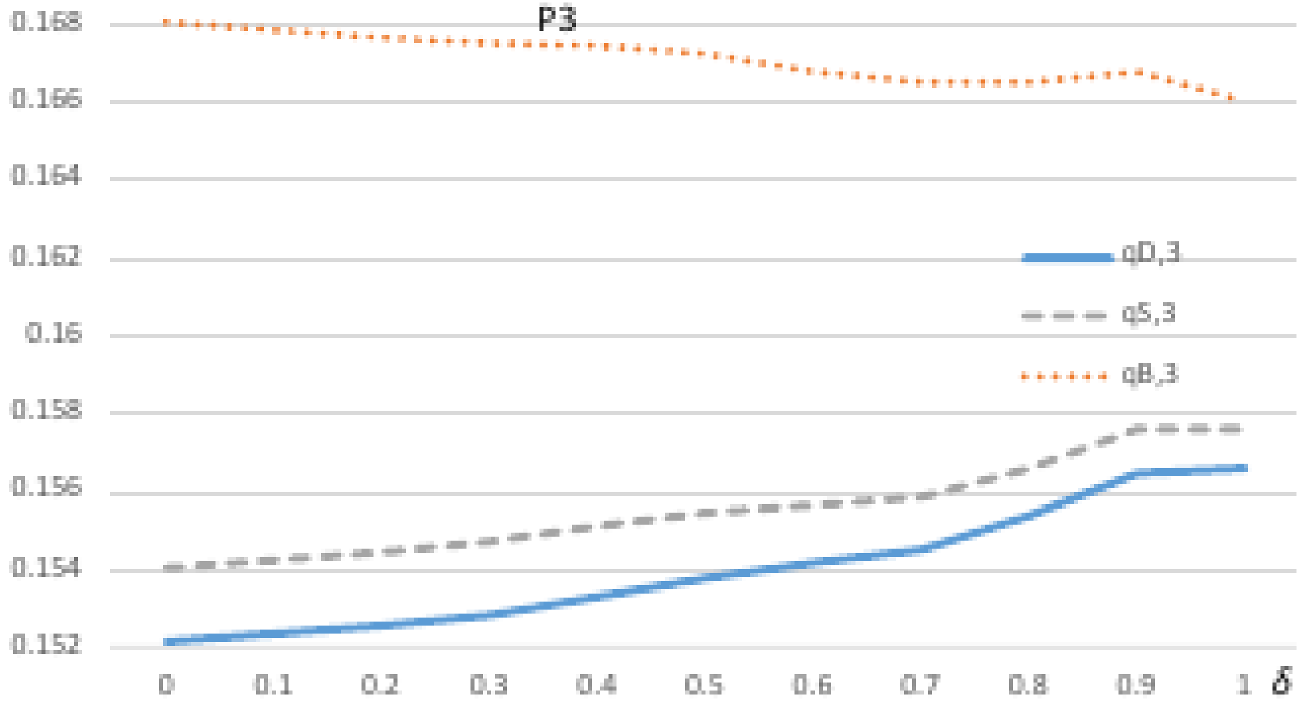

## Quarterfinal Round: Player 1

Decreasing tournament win probability due to less dominance as matrix irregularity increases. Since Player 1 can never lose its top ranking while it remains in the tournament, the static and dynamic rankings give identical outcomes.



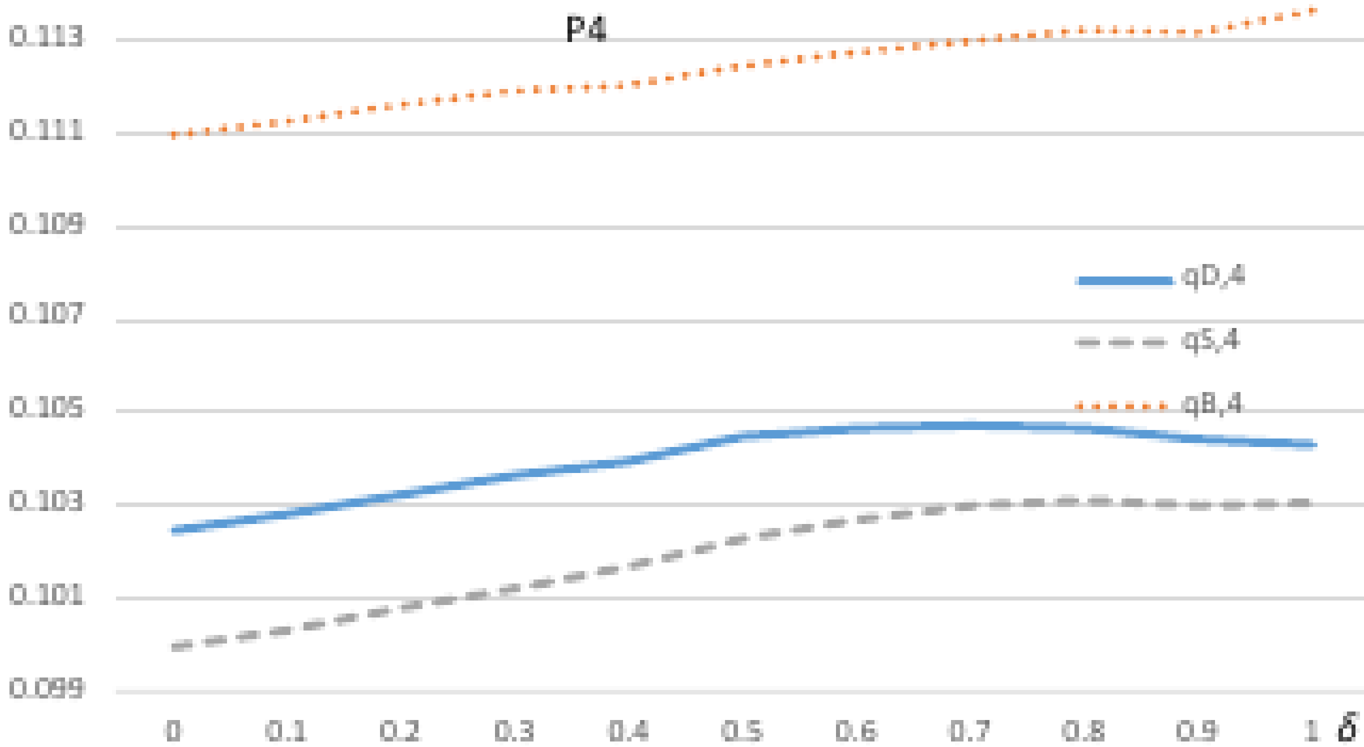

## Quarterfinal Round: Player 2

Decreasing tournament win probability due to less dominance as matrix irregularity increases. If Player 1 loses at the quarterfinal, 2 chooses first under the static ranking but not under the dynamic ranking.



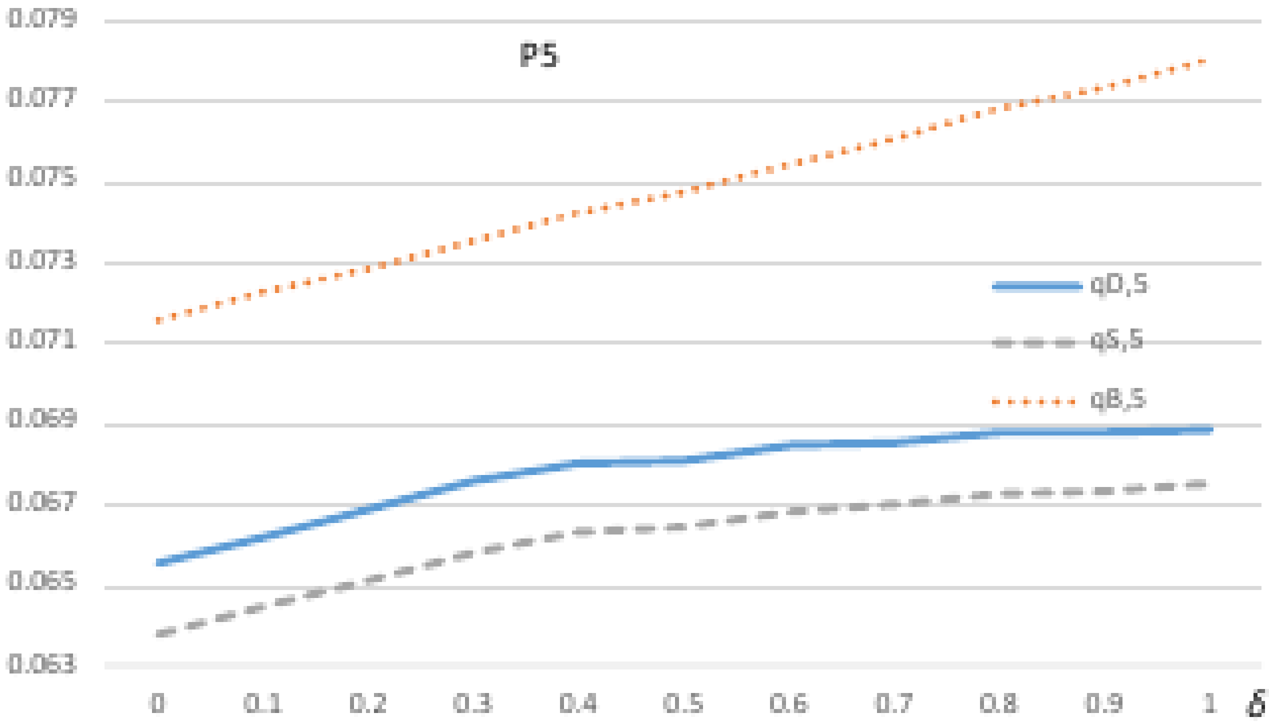

**Quarterfinal Round: Player 3**

Increasing tournament win probability due to less dominance by Players 1 and 2 as matrix irregularity increases. More benefit comes from retaining high original ranking under the static ranking design.



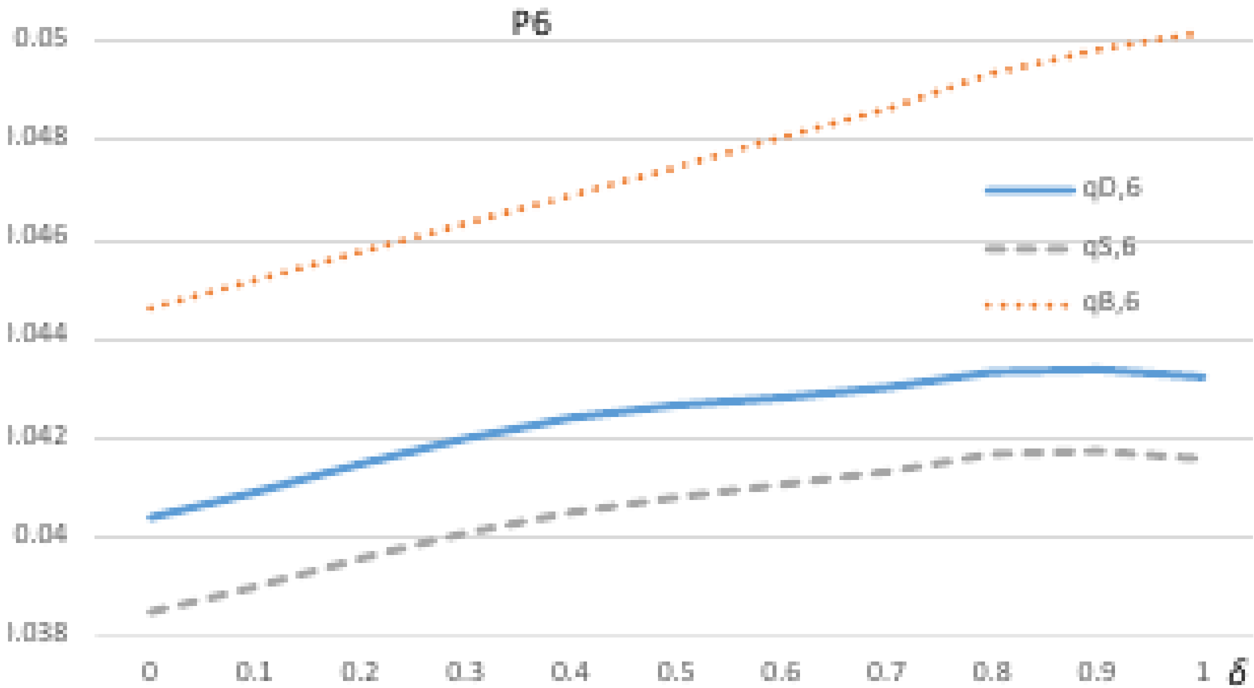

## Quarterfinal Round: Player 4

Unlikely to choose at the quarterfinal round, so the main benefit comes from increasing their ranking dynamically.



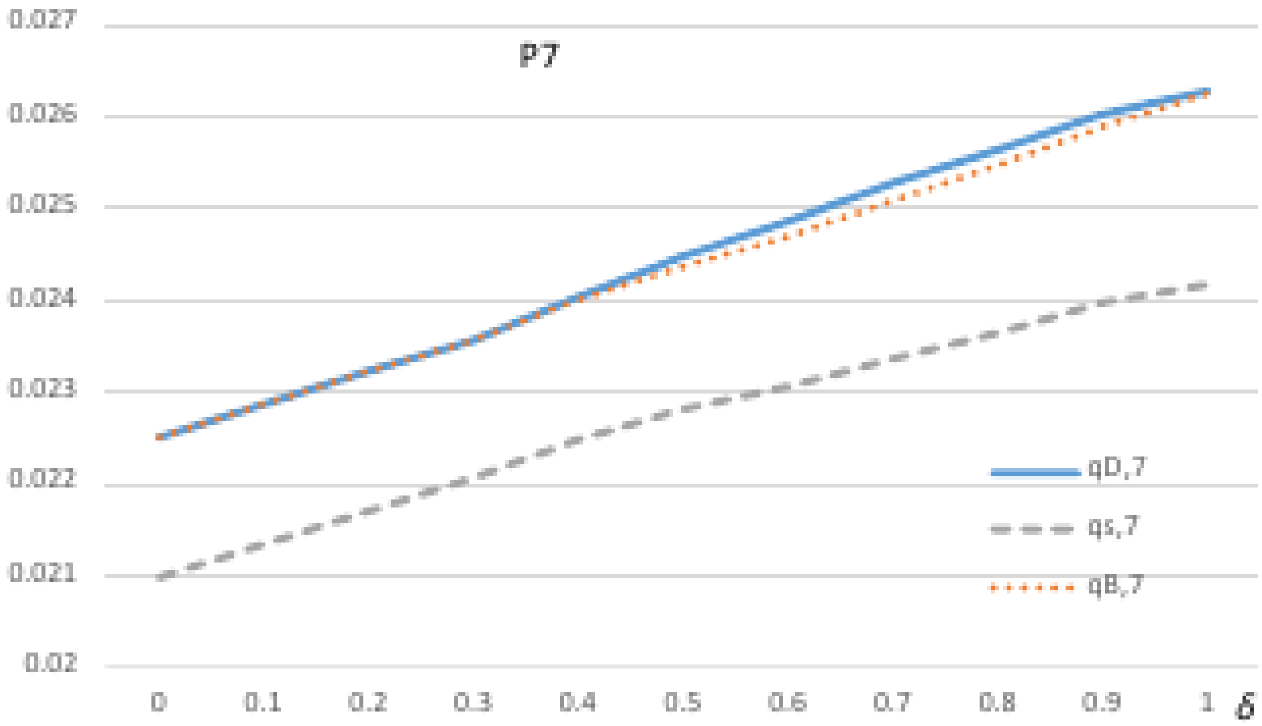

**Quarterfinal Round: Player 5**

For moderate matrix irregularity, it benefits from the chance to increase its ranking. But for high irregularity, the limited size of its choice set is outweighed by the increased probability of being chosen by a higher ranked player.



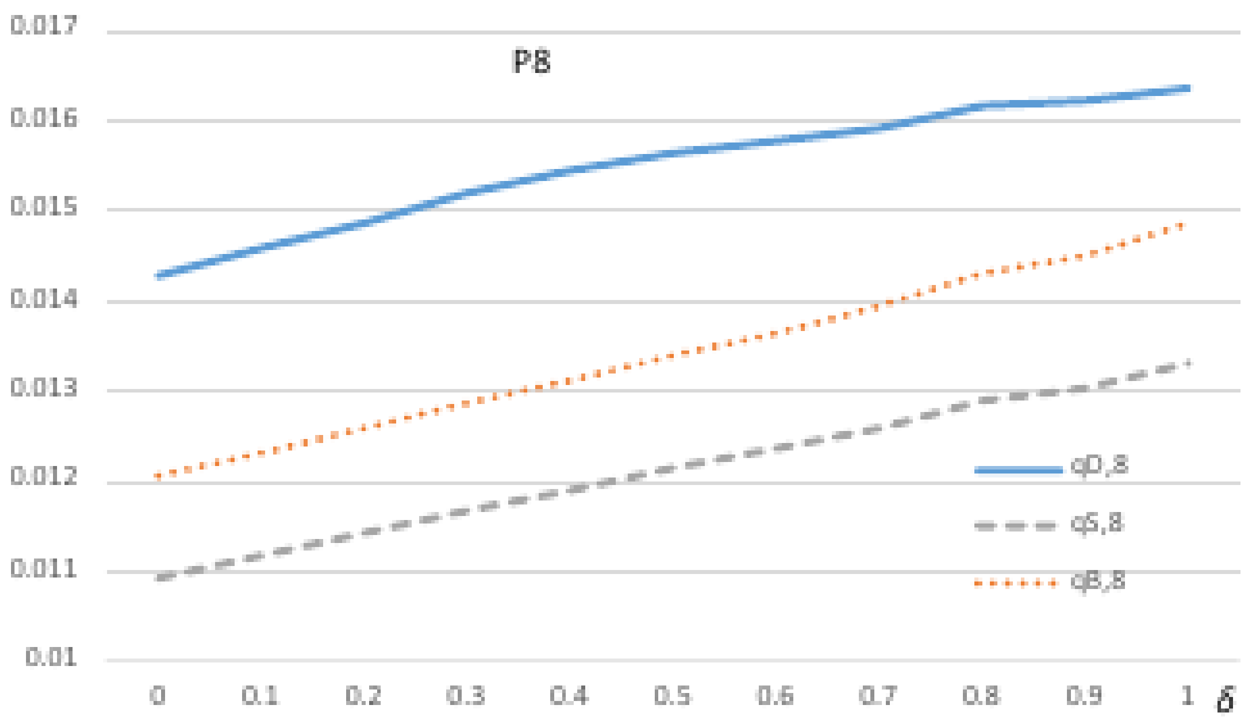

## Quarterfinal Round: Player 6

Increasing tournament win probability due to less dominance by other players. Never chooses under the static ranking, but may do so under the dynamic ranking.




# Quarterfinal Round: Player 7

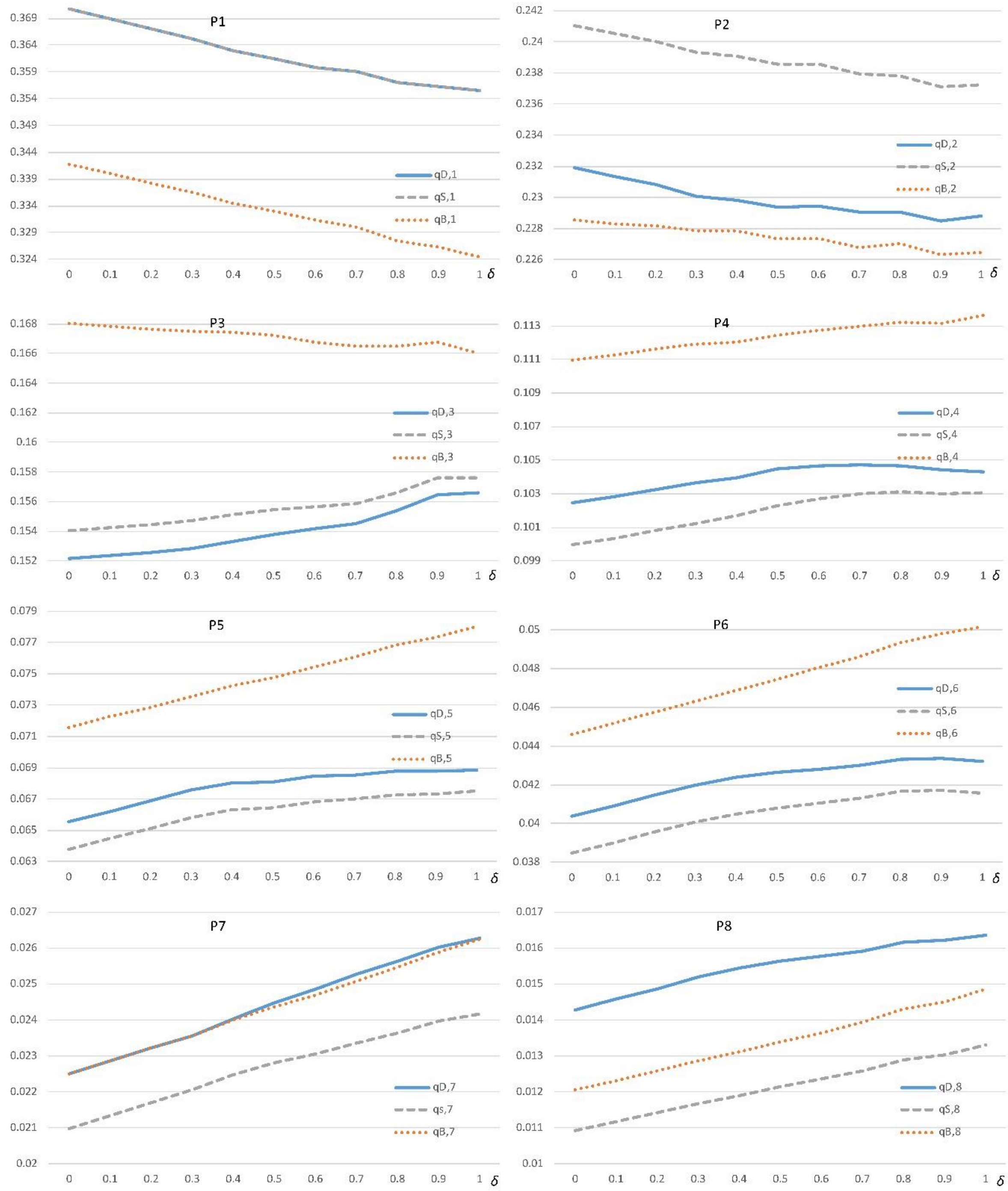
Increasing tournament win probability due to less dominance by other players. Never chooses under the static ranking, but may do so under the dynamic ranking.




**Quarterfinal Round: Player 8**

Increasing tournament win probability due to less dominance by other players. Never chooses under the static ranking, but may do so under the dynamic ranking.

# Quarterfinal Round Results with Full Information



# Probability that the Top Two Players Meet

Semifinal

Quarterfinal

| Parameter | $q_{D4,12}$ | $q'_{D4,12}$ | $q_{B4,12}$ | $q_{D8,12}$ | $q'_{D8,12}$ | $q_{S8,12}$ | $q'_{S8,12}$ | $q_{B8,12}$ |
|---|---|---|---|---|---|---|---|---|
| $\delta = 0$ | 0.358 | 0.000 | 0.358 | 0.257 | 0.000 | 0.257 | 0.000 | 0.258 |
| $\delta = .01$ | 0.356 | 0.000 | 0.356 | 0.256 | 0.000 | 0.256 | 0.000 | 0.256 |
| $\delta = .02$ | 0.354 | 0.000 | 0.354 | 0.254 | 0.000 | 0.254 | 0.000 | 0.254 |
| $\delta = .03$ | 0.353 | 0.000 | 0.352 | 0.252 | 0.000 | 0.252 | 0.000 | 0.252 |
| $\delta = .04$ | 0.352 | 0.000 | 0.351 | 0.250 | 0.000 | 0.250 | 0.000 | 0.251 |
| $\delta = .05$ | 0.352 | 0.000 | 0.349 | 0.250 | 0.000 | 0.250 | 0.000 | 0.249 |
| $\delta = .06$ | 0.351 | 0.002 | 0.348 | 0.249 | 0.001 | 0.249 | 0.001 | 0.247 |
| $\delta = .07$ | 0.348 | 0.013 | 0.346 | 0.249 | 0.003 | 0.249 | 0.003 | 0.245 |
| $\delta = .08$ | 0.345 | 0.022 | 0.344 | 0.248 | 0.007 | 0.248 | 0.007 | 0.243 |
| $\delta = .09$ | 0.339 | 0.039 | 0.343 | 0.247 | 0.011 | 0.247 | 0.011 | 0.241 |
| $\delta = .10$ | 0.337 | 0.049 | 0.341 | 0.247 | 0.015 | 0.247 | 0.015 | 0.240 |
| Mean | 0.349 | 0.011 | 0.349 | 0.251 | 0.003 | 0.251 | 0.003 | 0.249 |

Probability of Player $P1$ Meeting Player $P2$.

In the semifinal round, when irregularity is very high ($\delta \geq .09$), the chance for these players to meet in the final under our design is slightly less, but is more than compensated by the probability that they meet in the semifinal round.

However, in the quarterfinal round, even with increased probability for players 1 and 2 to meet in the semifinal round, our design still offers greater probability for them to meet in the final than the bracket design does.

# Order Preservation

| Parameter | $\tau_{D4}$ | $\tau_{B4}$ | $\tau_{D8}$ | $\tau_{S8}$ | $\tau_{B8}$ |
|---|---|---|---|---|---|
| $\delta = .04$ | 0.998 | 0.998 | 1.000 | 1.000 | 1.000 |
| $\delta = .05$ | 0.990 | 0.989 | 0.999 | 0.999 | 1.000 |
| $\delta = .06$ | 0.970 | 0.970 | 0.996 | 0.997 | 0.999 |
| $\delta = .07$ | 0.945 | 0.947 | 0.991 | 0.991 | 0.996 |
| $\delta = .08$ | 0.921 | 0.924 | 0.983 | 0.983 | 0.990 |
| $\delta = .09$ | 0.894 | 0.900 | 0.974 | 0.974 | 0.983 |
| $\delta = .10$ | 0.870 | 0.876 | 0.964 | 0.964 | 0.973 |
| Mean | 0.941 | 0.943 | 0.987 | 0.987 | 0.991 |

Kendall Coefficient for Player Rankings.

Recall that order preservation is a desirable property of tournament design (Horen and Riezman 1985, Vu and Shoham 2011).

When $\delta \leq 0.03$, all the coefficients are 1.000 for these randomly generated win probability matrices.

With increasing irregularity $\delta$, our design achieves results that are only slightly less consistent with the players' ranking than the bracket design.

As irregularity increases, the player's original ranking becomes a less accurate measure of the players' relative strength.

# Partial Information Assumption

Next, we will study a situation where each player has

(a) complete and accurate information about its own win probabilities against every other player, and

(b) approximate information about win probabilities between all other pairs of players:

- if the true probability is between 0 and 0.4, it is estimated as 0,

- if the true probability is between 0.4 and 0.6, it is estimated as 0.5, and

- if the true probability is between 0.6 and 1.0, it is estimated as 1.0.

By studying this situation, we obtain an understanding of the *robustness* of our tournament design to inaccurate estimation of win probabilities.

# Quarterfinal Round Results with Partial Information

For players other than 2 and 7, their outcome falls between those for the bracket and the dynamic choice design.

Player 2 benefits from the less accurate choices of Player 1. Player 7 is less likely to be chosen by Player 1 than under full information.

# Shirking (1 of 5)

**Remark:** From Vong (2017), it is not possible to eliminate shirking entirely without imposing on the tournament design a severe restriction that may reduce the tournament's appeal to spectators.

Due to the above remark, we consider the *final game* at the end of the preliminary round. It is in exactly this situation that shirking is most likely (which includes the England – Belgium and Indonesia – Thailand games).

# Shirking (2 of 5)

First, we have a negative result.

**Theorem:** If the subset of players which continues to the following round is *dependent* on the result of the final game, then it is not possible to eliminate shirking.

**Proof:** By example. Suppose that by winning the final game, player *P* eliminates its opponent *O* but allows into the next round another player *E which always beats every continuing player including P*. Then, player *P* must avoid winning the game or lose the tournament.

# Shirking (3 of 5)

Suppose the *subset of players* which continues to the next round is *independent* of the result of the last preliminary round game. This situation still includes the England – Belgium and Indonesia – Thailand games.

**Individual Strategic Behavior (ISB)**

**Theorem ISB:**

a. If exactly one fixed player *will continue* to the next round, our tournament design eliminates shirking by either player at the last preliminary round game.

b. If both players *will continue* to the next round, then our tournament design eliminates shirking by either player at the last preliminary round game.

**Proof:** A *higher ranking* at the next round gives any player who will continue a superset of choices against a fixed set of opponents with a fixed ranking - and therefore a fixed sequence of opponent choices - relative to a *lower ranking*.

In both cases, our tournament design eliminates shirking that occurs under a conventional bracket design.

# Shirking (4 of 5)

**Group Strategic Behavior (GSB)** [Thanks to Chen Chen and Alessandro Agnetis for raising this issue.]

**Remark:** All negative (pro-shirking) ISB results continue under GSB.

**Remark:** GSB is most likely to occur where a tournament contains multiple teams representing the same country, for example in tennis, table tennis, or badminton, at the Olympics.

**Theorem GSB [applying Theorem ISB to the GSB case]:**

(i). [Exactly one player $P$ continues.] The result in part (a) of the proof of Theorem ISB partly fails under GSB. It remains true that player $P$ will not shirk, but player $O$ may shirk to give player $P$ a higher ranking at the next round.

(ii). [Both players continue.] The result in part (b) of the proof of Theorem ISB partly fails under GSB. In a situation where one player's ranking will not change depending on the results of the game, but the other player's ranking will change, then the former player (but not the latter player) may shirk.

# Shirking (5 of 5)

For GSB, we have a further negative result, which applies to games where a draw is possible in preliminary round games (for example, World Cup soccer). The players are *P* and *O*.

**Theorem:** Consider the last game of the preliminary round in a tournament where a draw is possible. Then, there exists no design under GSB that prevents shirking by *both* players *P* and *O*.

**Proof:** By example. Consider a situation where only a draw between players *P* and *O* will allow both to continue, which is the only way to eliminate a third player *E* which always beats every continuing player including both P and O.

# Summary (1 of 2)

- To resolve several deficiencies of conventional tournament design, we propose a new design where players with a high ranking can choose their opponents at each round of the single elimination stage.
- This design is implemented for both static and dynamic rankings of players.
- We describe a dynamic programming algorithm that computes, for each player, the optimal sequence of opponents to choose and the resulting tournament win probability.
- This algorithm is computationally tractable up to the round-of-16.

# Summary (2 of 2)

- Using data from 1,902 men's professional tennis tournaments, we demonstrate the reasonableness of our tournament design.
- Our design allows flexibility for static and dynamic rankings, and the number of players who may choose their opponent(s).
- Compared to a conventional bracket design, our tournament design eliminates some elements of luck, provides reasonably increased probability for the top ranked player to win and also for the top two players to meet, and preserves ranking well.
- It also reduces shirking, and enables developing information to influence the tournament.
- An additional advantage of our design is increased fan interest from the unpredictability of later round matchups.

# Future Research

- Situations where a player knows its own win probabilities, but has *no information* about the win probabilities between any pair of other players. Here, a player may use a myopic strategy: choose an opponent which it can beat with highest probability. Our design using such a strategy still reduces shirking and is easier to analyze.

- We provide a related result.

# Without Any Knowledge of Games Between Other Players

**Theorem:** For any instance of an $n$-stage tournament, let Players $2, \ldots, 2^n$ be ordered such that $p_{12} \leq p_{13} \leq \cdots \leq p_{12^n}$.
Then, if player 1 chooses the lowest ranked available player at each round, this strategy guarantees that $q_1 \geq \prod_{i=1}^{n} p_{12^i}$, and this bound is tight.

**Proof:** We prove the lower bound by induction on the number of rounds played. Observe that, at round $k+1$, whatever previous results have occurred there must exist at least one remaining player with index of at least $2^{n-k}$.
Player 1 chooses that player as its opponent.

To show that the bound is tight, let $p_{ij} = 1$, for $2 \leq i < j \leq 2^n$.
When $k$ rounds remain to be played, the remaining $2^k$ players are exactly $1, \ldots, 2^k$.
Therefore, there is no possibility for Player 1 to achieve a higher value of $q_1$.

# Future Research

- Situations where a player knows its own win probabilities, but has *no information* about the win probabilities between any pair of other players. Here, a player may use a myopic strategy: choose an opponent which it can beat with highest probability. Our design using such a strategy still reduces shirking and is easier to analyze.
- Explore the application of our proposed tournament design to empirically-based studies of various sports and competitions, for example table tennis and chess.
- Study how the results of a particular round could be used to modify the win probability matrix, and consequently the choices of the players at later rounds.
- Allow players to have a different objective, for example maximizing the probability of reaching a particular round.
- Allow the ranking of the players to be adjusted dynamically, based on detailed performance within the tournament, as measured for example by margin of victory.

# Choosing Opponents in Tournaments

Nicholas G. Hall *

Zhixin Liu **

* The Ohio State University

** University of Michigan - Dearborn

Working Paper, Fisher College of Business, OSU

Are there any questions?

Are there any questions?
Are there any questions?
Are there any questions?
Are there any questions?
Are there any questions?
Are there any questions?