How an Economist Enjoys the NCAA Tournament

Bill Nelson | March 16, 2021

When filling in your NCAA bracket, is it always the best strategy to fill in the second round by picking the higher-ranked team in each matchup and then continue? By "best," I mean the strategy that yields the highest expected number of wins. And I'm assuming the rankings are correct in the sense that a higher-ranked team always has better than a 50 percent chance of beating a lower-ranked team.

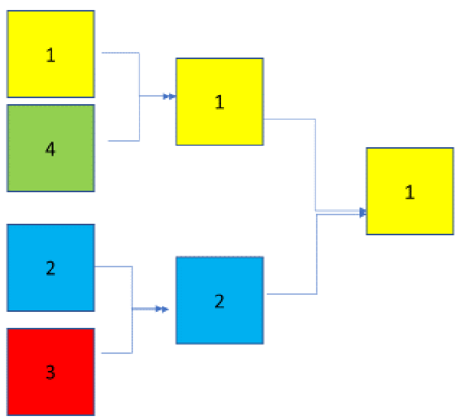
No, it's not always the best strategy, but it is if you make some stronger assumptions. Before going on, if you said, "No, it's not, because there are always upsets," please attend a probability and statistics class before reading this.

I'm going to deal just with a simple single-elimination tournament that starts with four teams. I'd be surprised if the results don't extend to bigger tournaments, but I am often surprised.

For simplicity, I will refer to teams by their rankings, and I define the probability of team *i* beating team *j* as w_{ij}. The rankings are correct in the sense that w_{ij} is greater than 50 percent whenever *i* is lower than *j* (that is, team *i* is ranked higher than team *j*).

The bracket that fills in the higher-ranked team is:

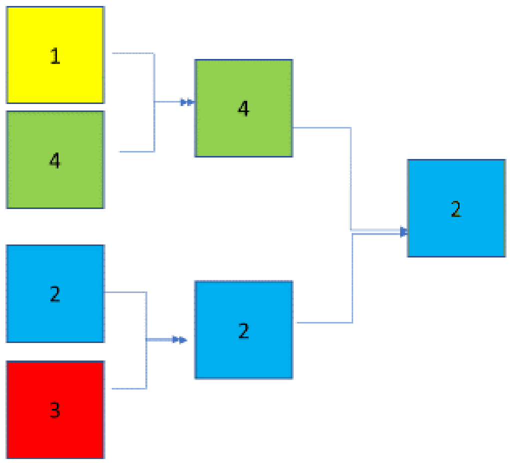
JUST PICKING WINNERS



To see that there can be cases when this isn't the best choice, consider the situation where, even though the rankings are correct, all the teams are quite evenly matched in the sense that the probability of any team beating any other team is close to 50 percent. In that case, it doesn't matter much how you fill the bracket in – the expected number of wins is about the same.

But suppose team 2 is likely to crush team 4; that is, w_{24} is almost equal to one. In that case, a better bracket choice would be:

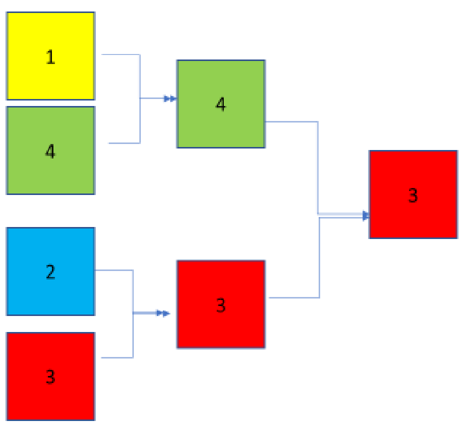
CASE 1: TEAM 2 DOMINATES TEAM 4



Why? Switching to picking team 4 over team 1 in the first round only reduces expected wins by a tiny amount but picking 2 over 4 (a matchup that is just about as likely as any other to occur) increases your expected wins by roughly 12½ percent of a win (50 percent times the 25 percent chance that the team 4 – team 2 matchup occurs. Intuitively, if roughly any matchup can occur, you might as well write down the one with the nearly certain outcome.

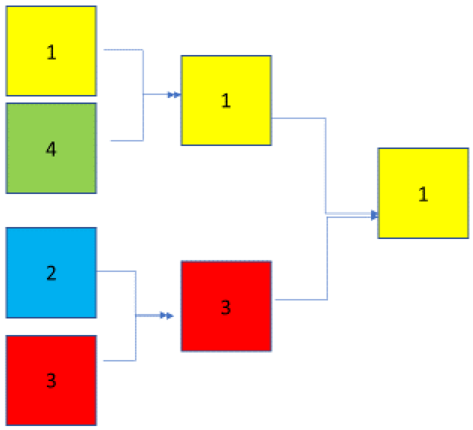
You get essentially the same result if team 3 would crush team 4. In that case, you should pick expected losers for both games in the first round:

CASE 2: TEAM 3 DOMINATES TEAM 4



You might think that a more intuitive case (case 3 below) would be when team 1 dominates team 3 with the implication that you should pick team 3 over team 2 in the first round. In that case, though, since you would still be picking the same final round winner (team 1), you gain no advantage and reduce your expected score a bit by picking team 3 in the first round.

CASE 3: TEAM 1 DOMINATES TEAM 3 (THIS IS NOT A BETTER BRACKET)



WHAT'S GOING ON HERE?

Returning to Case 1 – team 2 dominates team 4 -- the result is driven by the somewhat counterintuitive assumption that even though team 1 beats team 2, the probability that team 1 beats team 4 (about 50 percent) is lower than the probability that team 2 beats team 4 (nearly 100 percent). In Case 2 – team 3 dominates team 4 – even though team 2 beats team 3, the probability that team 2 beats team 4 (about 50 percent) is lower than the probability that team 3 beats team 4 (nearly 100 percent).

WHEN DOES IT JUST MAKE SENSE TO JUST PICK THE WINNERS?

We can rule out these cases by making a stronger assumption, I'll call it the "transitive property," about the rankings than simply that a higher-ranked team is expected to beat a lower-ranked team. Specifically, assume that the probability team *i* beats team *j* is higher than the probability that team *i* beats team k whenever j>k. For example, the probability of a number 1 seed beating a number 3 seed should be higher than the probability of a number 1 seed beating a number 2 seed.

As shown in the appendix, a corollary of the transitive property is that w_{ij} is greater than w_{kj} whenever *i* is less than k. That is, the probability a number 1 seed beats a number 3 seed is higher than the probability a number 2 seed beats a number 3 seed.

Going back to Case 1 – team 2 dominates team 4 – the assumption all the games are close matches except for the matchup between team 2 and team 4 violates the corollary of the transitive property. If the transitive property holds, then if the probability that team 2 beats team 4 is close to 100 percent, the probability of team 1 beating team 4 must be even closer to 100 percent. In that case, assuming team 4 beats team 1 in the first round is costlier in terms of expected wins than the benefit gained by advancing team 4 to the second round simply so that you can pick team 2. Case 2 is similar.

APPENDIX: PROOF OF THE COROLLARY

Transitive property: the probability team *i* beats team *j* is higher than the probability that team *i* beats team k whenever team k is ranked more highly than team *j*. That is, $w_{ij} > w_{ik}$ whenever *j*>k.

Narrowing the scope to just 3 teams, the transitive property says that

1. $w_{13} > w_{12} > w_{11}$ = 50 percent
2. $w_{23} > w_{22}$ = 50 percent.

Corollary: the probability that team *i* beats team *j* is greater than the probability that team k beats team *j* whenever team *i* is ranked more highly than team k. That is, w_{ij} is greater than w_{kj} whenever *i* is less than k.

In the case of 3 teams, the corollary says $w_{13} > w_{23}$.

Proof

From the transitive property

$W_{32} > w_{31}$

Multiply both sides by negative 1 (which switches the greater than to a less than)

$-w_{32} < -w_{31}$.

Add 1 to both sides

$(1-w_{32}) < (1-w_{31})$.

But $(1-w_{ij})$ is the probability that team *i* loses to team *j*, which equals the probability that team *j* beats team *i*, so

$w_{23} < w_{13}$.

QED

Disclaimer: The views expressed do not necessarily reflect those of the Bank Policy Institute's member banks, and are not intended to be, and should not be construed as, legal advice of any kind.