

NHL realignment effects on conference gap

Kilbourne Charrier, Michael Bauers, Clinton, Sam Downs

Motivation

Introduction

Methodology

Basic Idea

The basic idea behind the model is that all of the teams have a “skill level” that they play at or near during any given game. Each game is competitive because we aren’t sure how well either team will play, and even a team with a higher skill level can be beaten by another team for any number of reasons. We keep track of this skill level and draw other metrics based off of it, then use the metrics to compute a team’s score for the season. This score is used to calculate the team’s seed in their respective conference and division.

Creating the Model

To start out, we first need to make sure that we calculate the game-to-game performance levels given by τ and the standard deviation of the team’s talent levels given by σ . These are related to the following ratio, where p is the probability that the worst team in the league beats the best team in the league.

$$-\frac{2.8854}{\Phi^{-1}(p)} = \frac{\tau}{\sigma}$$

Based on the data from all of the seasons with 30 teams, we can calculate that the team with the worst record in the league beat the team with the best team about a quarter of the time. We choose to set p equal to 0.25 for this reason, making the ratio $\tau/\sigma = 4.28$. This means that there are four times the variability in the team’s performances in game than there is in the talent levels of the teams. The actual values of τ and σ don’t matter, just the ratio, but for simplicity we will set τ equal to 4.28 and σ equal to 1.

We then want to simulate a skill level metric for each of the thirty teams. We create the metric from the following distribution, with team i being equal to μ_i .

$$\mu_i \sim N(0, \sigma^2)$$

We draw the values from a normal distribution because in sports the overall performance of the teams is generally normally distributed. That is, there are a few really good teams, a few really bad teams, and a lot of teams in between.

We hold the team skill levels constant throughout the season, this make intuitive sense, because there are very seldom large changes to the team’s rosters or strategies throughout a season, but after each season teams often make at least some change.

For each game that is played in the season we draw a performance metric for both competing teams. This metric is drawn from a normal distribution with the mean being equal to the team’s skill level and the standard deviation being equal to τ^2 . This is based on the idea that a team with any skill level will play somewhere around their skill level, rarely playing much worse and rarely playing much better. This metric is calculated as such, with i indicating the home team and j indicating the away team.

$$\gamma_i \sim N(\mu_i, \tau^2)$$

$$\gamma_j \sim N(\mu_j, \tau^2)$$

Logically, the winner of each game would be decided by the higher value of γ , but in hockey it is possible for the game to go into overtime. This happens in about 22.4% of the games that were played over the past 15 seasons, so we must solve the following equation for α :

$$Pr(|\gamma_i - \gamma_j| < \alpha) = 22.4\%$$

Solving this equation gives us a value of $\alpha \approx 1.769$.

If there is a game in which $|\gamma_i - \gamma_j| \geq \alpha$, the team with the higher γ value is awarded a win and is therefore given two points while the team with the lower γ value is given 0 points. For games in which $|\gamma_i - \gamma_j| < \alpha$, the game is considered a tie.

In the case of a tie we linearly rescale the γ values with the following formula:

$$\zeta = \frac{(\gamma_i - \gamma_j) - a}{b}$$

Where a is the mean and b is the maximum of all $\gamma_i - \gamma_j$ values from that season.

Using the resulting value of ζ we use a Bernoulli distribution to simulate a weighted coin flip, with a success meaning that the home team wins and a failure meaning the away team wins.

Winner $\sim \text{Bernoulli}(\zeta_{ij})$

The winning team receives two points while the losing team receives one point.

We simulate 10,000 seasons using these principles and simulate the old and new rulesets by going through the different division brackets depending on the rules. We then calculate the conference gap for both the old and the new ruleset. We also save the number of points earned by each seed to do analysis on later.

Additional model details

Our model does not account for the real-world skill of teams, but because we are not doing any inference on specific team performance depending on rule changes this is not relevant. Because we are interested in the difference in scores of the 8th seeds we actually don't even need any real-world teams and can instead replace them with placeholders or just numbers.

An alternative choice we could have made is that in the event of a tie we could have given either team a 50% chance of winning instead of linearly rescaling the values and using a weighted coin flip, but we don't know exactly how the outcome of NHL overtime games usually end up skewing – whether it be in favor of the better team or if it is distributed more randomly.

Results

References

- Pettigrew, Stephen. “How the West Will Be Won: Using Monte Carlo Simulations to Estimate the Effects of NHL Realignment.” *Journal of Quantitative Analysis in Sports*, vol. 10, no. 3, 1 Jan. 2014, 10.1515/jqas-2013-0125. Accessed 10 Dec. 2020.