Modeling Roster Construction

Nehal Chigurupati

March 2024

1 Hard Cap Models

1.1 Single Year Scope

1.1.1 Variables

Let $\{p_i\}$, i $\in \mathbb{Z}_n$ be a collection of n players either currently on the roster or available to be added to the roster.

For each i $\in \mathbb{Z}_n$, let c_i be the proportion of the salary cap (i.e. the cost) of adding player p_i to the roster. To gauge such a cost, I "ranked" players in the NBA by estimated plus-minus and found the average proportion of the salary cap paid out to players less than or equal to three ranks away from p_i.

To assess the points added by a given player, let e_i^o be p_i's estimated offensive rating divided by 100 to approximate their points added per possession. Similarly, let e_i^d be p_i estimated defensive rating divided by 100, representing the points given up by a player per-possession.

Let n_p^i represent p_i's average possessions per game, and let n_p be the team average possessions per game in the 2023-24 season.

1.1.2 Objective Function

The function I seek to maximize is the ordinary Pythagorean win percent function, with exponent k=13.91. For p $\in \mathbb{Z}_2^n$ (i.e. a vector p where p_i is 1 if player p_i is in the roster, 0 if not), we have the following objective function:

$$\text{F}: \mathbb{Z}_2^n \rightarrow [0, 1]$$
$$\text{F(p)} = \frac{\sum_{i=0}^{n}[e_i^0 n_i^p p_i]^k}{\sum_{i=0}^{n}[e_i^0 n_i^p p_i]^k + \sum_{i=0}^{n}[e_i^d n_i^p p_i]^k}$$

The $[e_i^0 n_i^p p_i]$ term represents the points added by a given player, given by their estimated points per possession, e_i^0, times their average number of possessions per game, n_i^p, times p_i to ensure that the term disappears when the player is not included in the roster. Similarly, the $[e_i^d n_i^p p_i]$ term represents points given up, with e_i^d representing the points given up per possession by player p_i.

1.1.3 Rostered Players Constraint

Let J $\subseteq \mathbb{Z}_n$ represent the players already on the roster of the team being augmented. For instance, if j $\in \mathbb{J}$, then player p_j is on the team's roster, and cannot be replaced (i.e. they are a mandatory part of any salary cap or expected win percent calculations).

To ensure the optimization respects rostered players, I introduced the following constraint, requiring any p in the parameter space $\mathbb{Z}_2^n$ satisfy:

$$p_i = 1 \text{ if } i \in \mathbb{J}$$

1.1.4 Salary cap constraint

In this model, the salary cap is modeled as a hard limit. This limit is not necessarily equal to the actual 2023-24 salary cap. Rather, it is represented as a percentage of the cap, allowing for teams to spend up to 2x the cap. Although simplistic, in practice, this model of the CBA's salary system may be sufficient, given that GMs face caps imposed by ownership related to their willingness to enter the luxury tax.

Given p_i and c_i as above, as well as some t $\in$ [.95, 2], I imposed the following constraint on the optimization:

$$\sum_{i=0}^{n} c_i p_i \leq t$$

For instance, if t=1.5, then the maximum salary able to be paid out is equal to 1.5 times the 2023-24 salary cap.

1.1.5 Roster size constraint

The above problem considers only players on a uniform contract. To account for the fact that two-way and 10-day players could fill slots on rosters, some variability in the size of the team roster is permitted. The following constraint is imposed, requiring that the roster include at least 12 and no more than 15 players:

$$12 \leq \sum_{i=0}^{n} p_i \leq 15$$