Modeling Roster Construction (v3)

Nehal Chigurupati

March 2024

1 Informal Description

This model seeks to answer the following question:

Given a set of constraints derived from financial considerations, league rules, player demands and current contracts, what roster yields the highest expected regular season win percentage?

To answer this question, for an arbitrary team, I took a set of players who are eligible to be added to the roster (2024 free agents, in particular, but the choice is arbitrary), the set of players who are currently on the team roster who are not on contract for the next season, and the set of players who are currently on the team roster and are on contract for the following season. I then approximated the percent of the salary cap each of these players would likely demand (for players on contract for next season, this quantity was equal to the base annual compensation dictated by their existing contract), and tried to craft a roster that yields the highest expected regular season win percentage. I required that any roster satisfy the following constraints:

- **(Salary cap constraint)** The sum of the annual salaries demanded by each player on the roster (minus any salaries exempt from cap calculations by the CBA) must be less than the 2023-24 salary cap.
- **(Minimum salary constraint)** The sum of the annual salaries demanded by each player on the roster (minus any salaries exempt from cap calculations by the CBA) must be greater than or equal to 90 percent of the 2023-24 salary cap.
- **(Total budget constraint)** The sum of the annual salaries demanded by each player on the roster (including those exempt from cap calculations) must be less than or equal to some predefined amount.
- **(Roster size constraint)** The total number of players on a roster must be no greater than 15 and no less than 12. I allowed the size of the roster to shrink to as few as 12 players to account for two-way and 10-day deals, given that my model only considers players seeking to sign a uniform contract.

- **(Rostered players constraint)** All players on contract for the following season must be included in any roster configuration.
- **(Play time constraint)** The sum of the average possessions per game for each player on the roster must not exceed 5 times the team league-average possessions per game, ensuring that the team does not over promise playing time.

  The current iteration of the model computes the salary cap constraint with consideration for Bird, Early Bird, Non-Bird, and Minimum Salary exceptions. I'm working on expanding support.

2 Variables

2.1 Parameter Space

Let $\{p_i\}$, i $\in \mathbb{Z}_n$, $p_i \in \mathbb{Z}_2$ be a collection of n players either currently on the team roster or available to be added to the roster. p_i attains the value 1 when player i is included in a roster configuration, 0 otherwise. Let J, I $\subseteq \mathbb{Z}_n$ disjoint correspond, respectively, to the indices of the players from the 2023-24 season who are returning the next season and the players from the 2023-24 season who will become free agents in the off season. Let ϕ_J: $\mathbb{Z}_n \to \mathbb{Z}_2$ be an indicator function, evaluating to 1 if the input index is located in J (i.e. if player i is currently rostered and will be returning the following season). Similarly, let ϕ_I: $\mathbb{Z}_n \to \mathbb{Z}_2$ evaluate to 1 if the input index corresponds to a player who is currently rostered but will be a free agent in the off season, 0 otherwise.

The parameter space, without any refinements induced by constraints, corresponds to the following set:

$$X := \{p \in \mathbb{Z}_2^n\colon\ p_i \geq \phi_J(i),\ i \in \mathbb{Z}_n\}$$

2.2 Player Costs

Consider some player p_i. Let c: $\mathbb{Z}_n \to \mathbb{R}$ map player i to their annual "cost". Let s_i denote the annual salary of p_i provided by their latest contract. Let e_{pm}: $\mathbb{Z}_n \to \mathbb{Z}_n$ map player i to their rank in a list of players p_i, sorted by estimated plus-minus (greatest to least). We can force e_{pm} to be bijective by arbitrarily settling any ties that appear. With a bijective e_{pm}, we can then induce an inverse e_{pm}^{-1}: $\mathbb{Z}_n \to \mathbb{Z}_n$ that maps estimated plus-minus ranks to the index of the player that resides at the rank.

Now, we are ready to create a closed form expression for the cost of player i, c(i).

$$c(i) := s_i\phi_J(i) + \max\left\{ s_i,\ \frac{\prod_{k=min\{n,e_{pm}(i)+3\}}^{max\{0,e_{pm}(i)-3\}} s_{e_{pm}^{-1}(k)}}{min\{n,e_{pm}(i)+3\}-max\{0,e_{pm}(i)-3\}} \right\} [1 - \phi_J(i)]$$

Breaking down this expression, consider an arbitrary player p_i. If player i was on the team roster in the 2023-24 season and is on contract for the following season (i.e. ϕ_J(i) = 1), then their cost corresponds to s_i, the annual salary guaranteed by their existing contract.

On the other hand, if player i is not on contract for the next season, then their annual "cost" corresponds to the maximum of two quantities:

- The annual salary guaranteed by p_i's last contract
- The average annual salary earned by players who are similarly ranked by estimated plus-minus to p_i (three closest in both directions)

2.3 Player Skills

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

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

2.4 Player Veteran Status

Let V: $\mathbb{Z}_n \to \mathbb{Z}$ map a player index i to the number of consecutive years, up to and including the 2023-24 season, that player p_i has been on the roster of an NBA team.

2.5 Player Exception Eligibility

In my current iteration of the model, I've accounted for the following four salary cap exceptions:

- Bird Rights
- Early Bird Rights
- Non-Bird Rights
- Minimum Salary Exception

p_i's salary is eligible for one of the Bird exceptions if i $\in$ I, i.e. if the player has been on the team roster for one or more seasons and is now negotiating a new contract.

To handle eligibility for the minimum salary exception, let H: $\mathbb{Z} \rightarrow \mathbb{R}$ send a given number of years of service in the NBA to the corresponding minimum salary dictated by the CBA. For the 2024 season, H is determined by the following:

- H(0) (i.e. rookie) = 953,000
- H(1) = 1.53 million
- H(2) = 1.71 million
- H(3) = 1.79 million
- H(4) = 1.84 million
- H(5) 1.99 million
- H(6) = 2.15 million
- H(7) = 2.30 million
- H(8) = 2.46 million
- H(9) = 2.47 million
- H(10+) = 2.72 million

Let K $\subseteq \mathbb{Z}_n$ denote the indices of players who qualify for the minimum salary exception, formally defined as:

$$\mathrm{K} := \{\mathrm{i} \in \mathbb{Z}_n\text{: } \mathrm{c(i)} \leq \alpha \mathrm{H} \cdot \mathrm{V(i))}\}$$

where $\alpha \in \mathbb{R}$, $\alpha > 0$. α is a multiplicative constant used to account for the fact that players may be willing to drop their demanded salary down to the minimum, canonically set to $\frac{3}{2}$. Let ϕ_K: $\mathbb{Z}_n \rightarrow \mathbb{Z}_2$ be the subset inclusion function, evaluating to 1 if i $\in$ K, 0 otherwise.

2.6 Total Budget

Let β denote the total budget for player salaries, including all salaries exempt from cap calculations. Let S denote the salary cap.

3 Objective

The function I maximize is the ordinary Pythagorean win percent function, with exponent k=13.91.

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

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.

4 Constraints

4.1 Total Budget

I imposed the following constraint to ensure that all rosters adhere to the budget limit β:

$$\sum_{i=0}^{n} \text{c(i)p}_i \leq \beta$$

4.2 Salary Cap

As before, let S denote the salary cap. The salary cap constraint is formalized as follows:

$$\sum_{i=0}^{n} \Big[\text{ c(i)[1-}\phi_I\text{(i)][1-}\phi_K\text{(i)]p}_i \Big]$$

The [1 - ϕ_I(i)] factor excludes players eligible for Bird, Early Bird, or Non-Bird rights from the cap calculation, while the [1-ϕ_K(i)] term excludes players eligible for the minimum salary exception from the cap calculation.

4.3 Roster Size

I require that rosters consist of no greater than 15 players and no fewer than 12. I allow for as few as twelve players to account for my model's lack of support for 10-day and two way contracts. Formally, this roster size constraint can be expressed as:

$$12 \leq \sum_{i=0}^{n} \text{p}_i \leq 15$$

4.4 Rostered Players

Recall the subset $J \subset \mathbb{Z}_n$, indicating the indices of players who are on contract for the next season. To force these players onto any roster, I introduced the following constraint:

$$\mathrm{p}_i \geq \phi_J(\mathrm{i})\ \forall \mathrm{i} \in \mathbb{Z}_n$$

4.5 Playing Time

I require that the sum of the average possessions per game of each player on the roster must be no greater than 5 times the league-average possessions per game for teams:

$$\sum_{i=0}^{n} \mathrm{n}_p^i \mathrm{p}_i \leq 5\mathrm{n}_p$$

5 Optimization

I performed the resulting binary, non-linear optimization using GEKKO for Python. To try it out, visit my website here.