

Exploratory Data Analysis

Group 6

Measuring Lost Potential

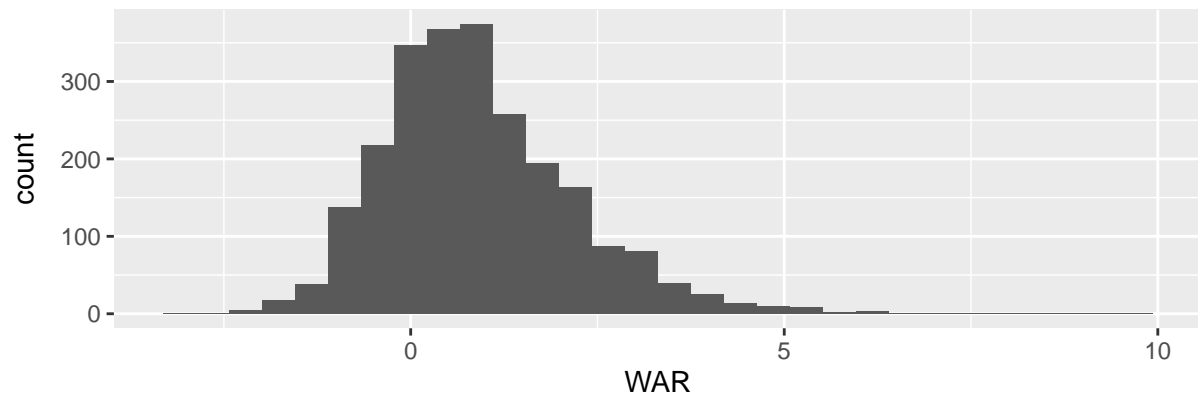
To reduce the inherently vague and amorphous concept of “hurting the game” into a measurably quantity, we fleshed out the “better players forced into retirement” idea from our proposal to define a specific type of player, the Couldabeen (CBn).

As previously laid out, a Couldabeen is a player entering early retirement whose performance in their last season of major league play suggests that they were replaced on the field by worse players. In order to qualify for Couldabeen classification, three things must be true of a player’s final season: a) the player was 35 or younger at midnight on June 30th; b) the player pitched more than 50 innings if a pitcher, or played in 30 games if a position player; and c) the player’s Wins Above Replacement stat for the season was within one standard deviation of the average rookie that year or better. Our data set includes all MLB seasons between 1969, when the MLB added four new franchises (thus greatly expanding the player pool), and 2018 (2019 and 2020 are excluded to prevent ambiguity over which players are actually retiring).

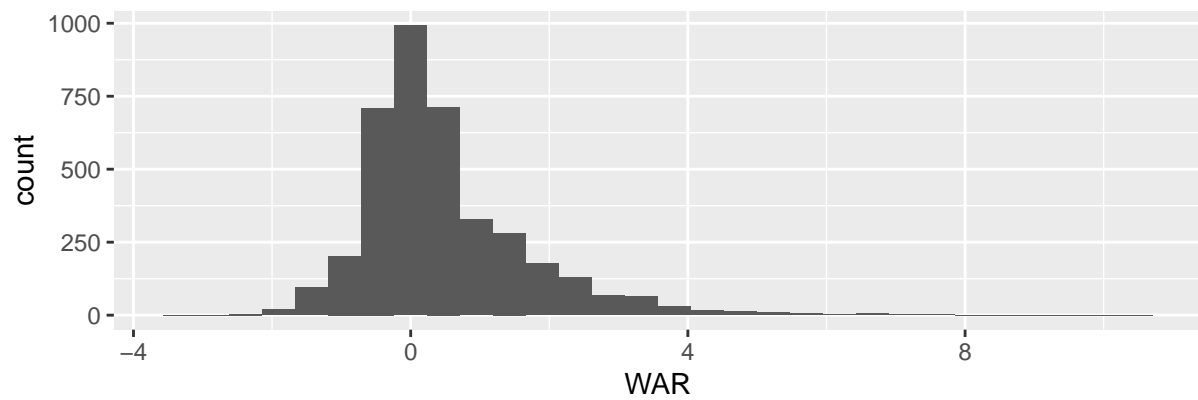
Histograms of Rookie WAR

To find the threshold defined in c), we compiled the season Wins Above Replacement stats from every rookie player between 1969 and 2018 who pitched more than 50 innings if a pitcher or played in 30 games if a position player. Although in theory they measure the same thing, the way WAR is calculated for pitchers is so different from WAR for position players that they may as well different statistics and are often listed as such. We therefore split the rookie data into two data sets, one for pitchers and one for everyone else.

Distribution of Pitcher WAR in Rookie Season Between 1969 and 2018



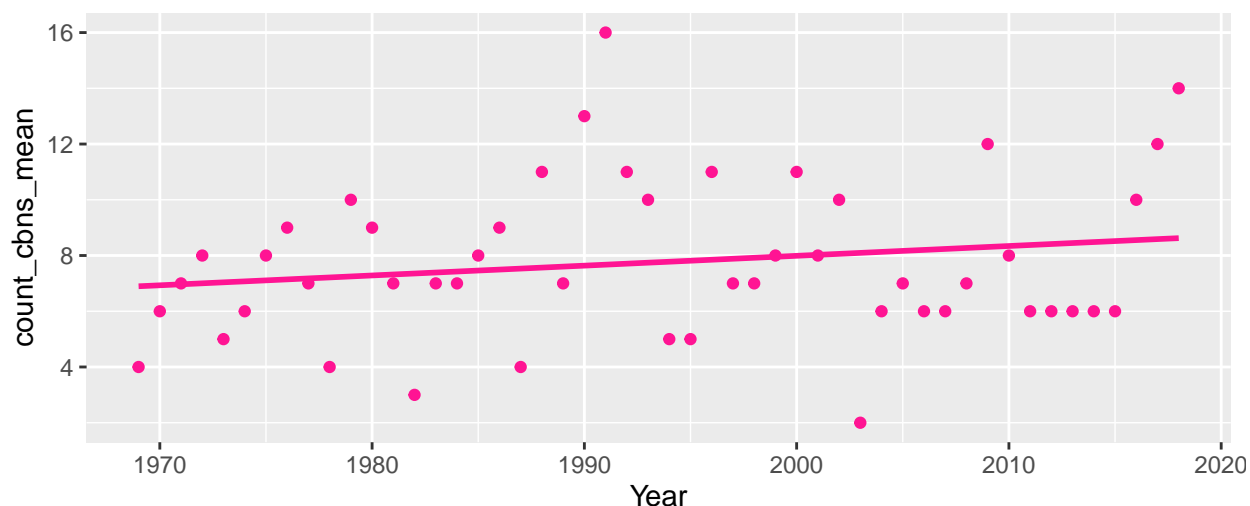
Distribution of Position Player WAR in Rookie Season Between 1969 and 2018



Counting Couldabeens

We then calculated our WAR thresholds for each season and counted the number of Couldabeens (pitcher and position player) retiring that year. We find two reassuring signs for our data set. First, an upward trend in the number of Couldabeens over time. Second, several important dates in MLB salary regulations correspond to visible Couldabeen spikes: the 1972 player strike, the first in history, which secured additional pension funding and players' right to salary arbitration; declining profits and salary cap proposals in the early 1990s and ensuing player strike, the *longest* in history; the reintroduction of the “luxury tax” in 2002; and the implementation of harsher luxury tax rules after the 2017 season.

Couldabeens: Pitchers



Couldabeens: Position Players

