Salary Inequality and Team Performance Across MLB, the NBA, and the NFL

Front Office: azimbali, aa146, ereed2, csaueres

Hypothesis

In baseball, a player only impacts the game when he is hitting (1/9 of the time) or when the ball is hit to him (infrequently). In football, most players impact a play infrequently (with perhaps a couple exceptions, such as quarterback). In basketball, however, it is possible for a single player to influence the game on every single offensive and defensive possession (it is not uncommon for elite players to either score or assist in over half of his team's points). We therefore hypothesize that salary inequality (spending exorbitantly on a few elite players at the expense of the rest of the roster) should be positively correlated with team performance in the NBA, but negatively correlated with team performance in MLB and the NFL.

Data

Data for MLB and the NBA is scraped from sportsreference.com (manually for MLB, and using an API for the NBA). Sportsreference.com is among the most reputable sites for sports data. NFL data is scraped from USA Today, which is a reputable news source that we trust provided accurate data. We collected MLB data from 1999 – 2018 (no salary data available for 2019), NBA data from 2000 – 2019, and NFL data for the years 2001, 2003, 2005, 2007, and 2009 (unfortunately, this is the only data that is available for free for the NFL). Data collected include player name, year, team they played for, player salary, and team record (wins and losses).

Findings

Claim #1: Gini coefficient (a measure of salary inequality) is related to win rate (a measure of team performance) in MLB and the NBA, but not in the NFL.

Support for Claim #1: Two hypothesis tests provide support for this claim. First, we group teams by win rate and concatenate team salary distributions within each group to find a joint salary distribution for high win rate teams and a joint salary distribution for low win rate teams. A Kolmogorov-Smirnov test (a non-parametric test for testing equality of distributions) finds that there is a statistically significant (5% significance threshold used throughout study) difference in salary distribution between high and low win rate teams in MLB and the NBA, but not in the NFL. This test does not directly look at the Gini coefficient, but it indicates that some relationship exists between win rate and salary distribution. To further examine this relationship, we compare Gini coefficients for low win rate teams to Gini coefficients for high win rate teams (without concatenating salary distributions). A two-sample T-test shows that there is a statistically significant difference in the means of the Gini coefficient between high win rate teams and low win rate teams in MLB and the NBA, but not in the NFL. This test provides evidence that Gini coefficient is related to win rate in MLB and the NBA, but does not say anything about **how** they are related.

Sport	KS-test (Salary Dist.) p-value	T-test (Gini Coef.) p-value
NBA (basketball)	8.211e-17***	1.903e-15***
MLB (baseball)	8.081e-13***	0.0005***
NFL (football)	0.156	0.918

Claim #2: Gini coefficient is positively correlated with win rate (more salary inequality is correlated with better team performance) in the NBA, negatively correlated with win rate in MLB, and not significantly correlated with win rate in the NFL (corroborating claim #1).

Support for Claim #2: In order to determine **how** Gini is related to win rate, we run a multiple linear regression with win rate as the dependent variable and with Gini coefficient and payroll as the independent variables. Payroll influences both Gini coefficient (higher payroll correlated with higher Gini) and win rate (higher payroll correlated with higher win rate), and so we control for this potentially confounding variable by including it in our multiple regression. Payroll is normalized using min-max scaling to adjust for changes in average total payroll over time (increases due to inflation).
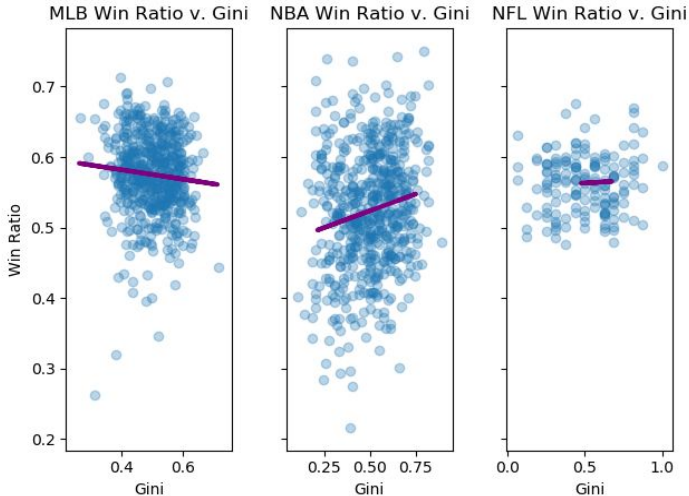
Sport	Regression Coefficient for Gini	P-value for Gini
NBA (basketball)	0.1665	0.021*
MLB (baseball)	-0.1173	0.016*
NFL (football)	-0.1473	0.719