Econ 150 Writeup

Connor McCollick, Griffin Holcombe

1. Introduction and Question

This study examines whether MLB team spending is associated with seasonal performance. The question: Does higher team payroll predict a higher win–loss ratio? The analysis evaluates three spending measures: total payroll, active 26-man payroll, and the ratio of active to total payroll. The objective is to determine whether spending levels meaningfully explain variation in team success.

2. Data Description and Sources

The dataset contains MLB team payroll information from 2011–2024, sourced from Kaggle (Christopher Treasure). All payroll figures were adjusted for inflation at 2.5% annually to convert values into 2024 dollars. The primary variables include total payroll, active 26-man payroll, the active-to-total payroll ratio, and each team's win–loss ratio. Scatterplots show weak to moderate positive associations between spending and performance, with active payroll appearing visually strongest. Data issues include inconsistent reporting across seasons and reliance on aggregated third-party payroll estimates. Cleaning involved removing incomplete records, computing inflation-adjusted series, and generating ratio variables. The scatterplots from models 1, 2, and 3 can be seen in the appendix.

3. Methodology

The analysis uses a generalized linear model estimated by ordinary least squares. The specification is $\text{WinLossRatio} = \beta_0 + \beta_1(\text{PayrollMeasure}) + \varepsilon$. The coefficient β_1 captures the marginal association between each payroll measure and team performance. The model assumes linear relationships, homoscedastic errors, and independent observations across teams and

seasons. These assumptions are unlikely to fully hold in practice because team performance depends on many omitted variables, and payroll effects may be nonlinear. As a result, the model's explanatory capacity is inherently limited.

4. Results and Analysis

Model 1 uses inflation-adjusted total payroll. The coefficient is approximately 1.802e-09, positive and statistically significant, but extremely small in magnitude. The model's R^2 is 0.115, indicating that total payroll explains 11.5% of the variation in win–loss ratios. Residuals show substantial dispersion, confirming weak predictive power. Model 2 uses inflation-adjusted active 26-man payroll, which performs better with an R^2 of 0.219. This measure has a stronger positive association with performance, though large residual variance remains. Model 3 evaluates the active-to-total payroll ratio and yields an R^2 of 0.186. The ratio correlates with success but is less predictive than absolute active payroll. Across all models, payroll variables show statistically detectable relationships with outcomes, but most variation remains unexplained due to factors unrelated to spending.

5. Conclusions

The results show that payroll levels in MLB are positively associated with success, with active payroll being the strongest predictor of team performance. However, even the best model explains only a modest share of win–loss variation. Financial investment contributes to success but does not dominate performance outcomes. This analysis answers the research question by showing that payroll matters, yet its impact is limited relative to other organizational and player-specific factors.

6. References

Kaggle. MLB Team Payrolls 2011–2024.

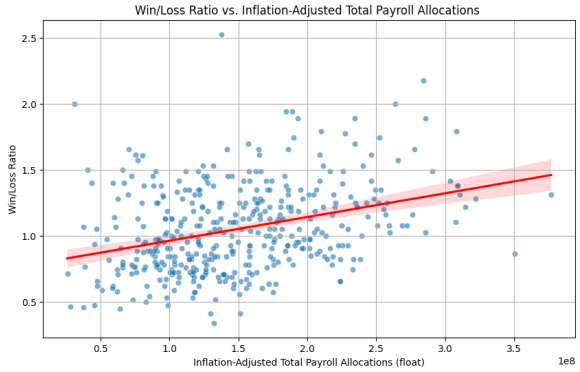
https://www.kaggle.com/datasets/christophertreasure/mlb-team-payrolls-2011-2024

Relevant empirical literature on sports economics and payroll–performance relationships.

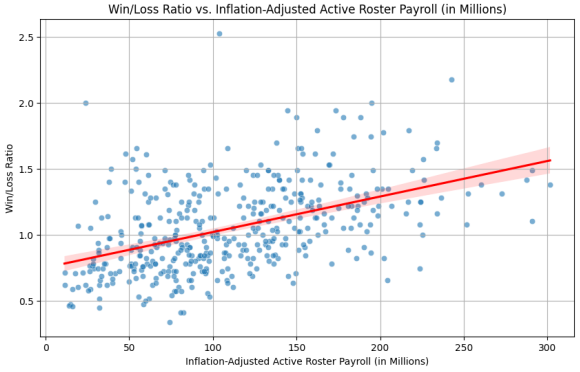
Google Gemini was used with Modeling and code assistance for the project.

7. Appendix

Model 1 Scatterplot:



Model 2 Scatterplot:



Model 3 Scatterplot: