

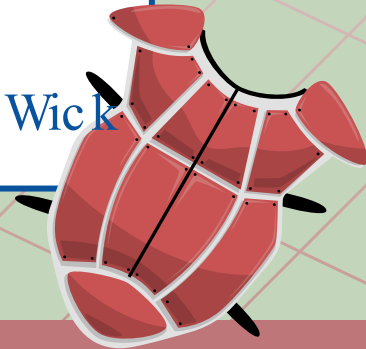
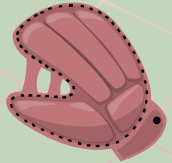
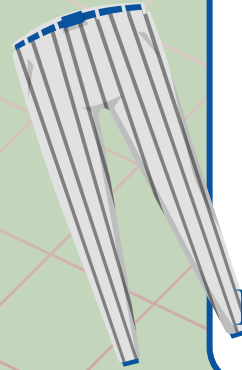


Wharton

UNIVERSITY of PENNSYLVANIA

Are “Surprise Teams” in the MLB in 2023 More Surprising?

By: Chad Federico, Kotaro Nagano, and Alejandro Wick



Introduction & Problem

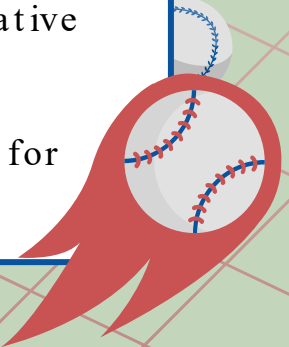
Many teams, about halfway through the 2023 season, are playing much better (Rays, Reds, Diamondbacks, Marlins) or worse (Mets, Padres) than expected, especially considering team payrolls. Injuries have also led to unexpected team performances on the field.

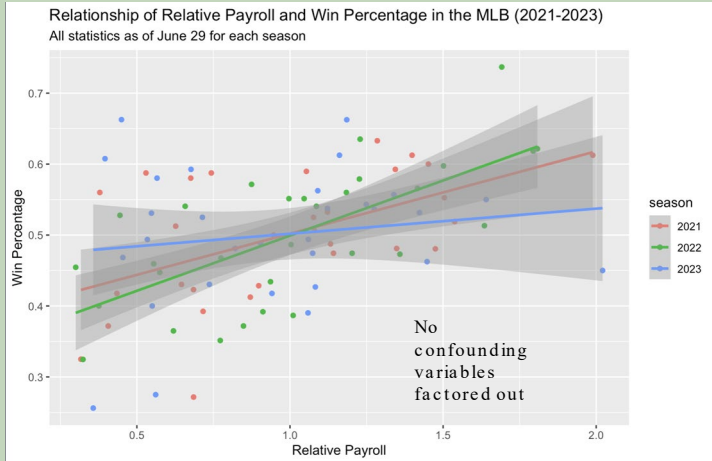
Surprise teams usually have a pythagorean win percentage (based on the total amount of runs scored and allowed) quite different than their actual win percentage.

Luck/chance is always a factor in “surprise teams” such as these, but is the extent to which we’re seeing these teams surprise is greater in 2023 than recent years?

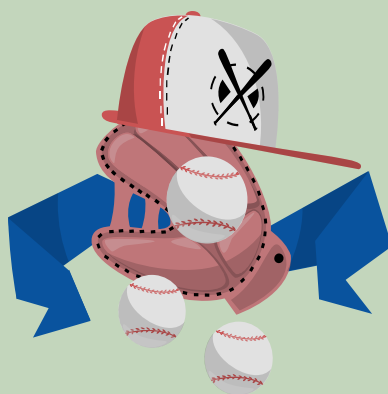


Procedure

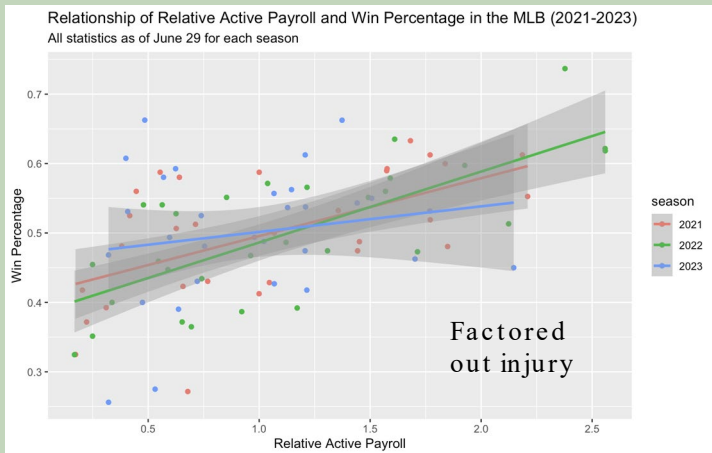
- Data from 2021-2023 MLB Seasons as of June 29th for each season (roughly 80 games or half of a season)
 - Found relative payrolls (full team payroll/ median team payroll) for each team in each season
 - To account for injuries, we also found relative active payrolls (active man payroll/ median active man payroll)
 - Slightly different amount of games played depending on the year, so we used win %
 - To account for “luck”, we also used the pythagorean win %
 - Plotted Raw and Pythagorean Win Percentages vs. Relative Payrolls and Relative Active Payrolls
 - Created a graph comparing the correlation coefficient for all 4 relationships for each year
- 



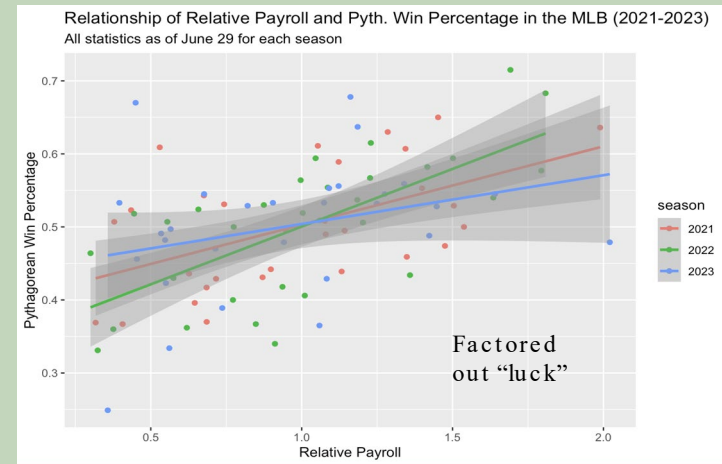
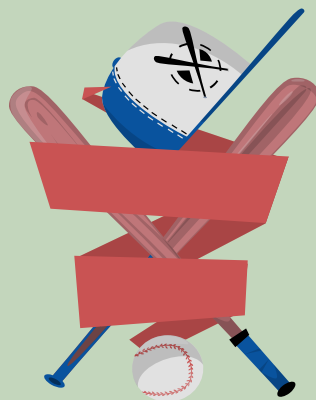
Win % vs Relative Payroll



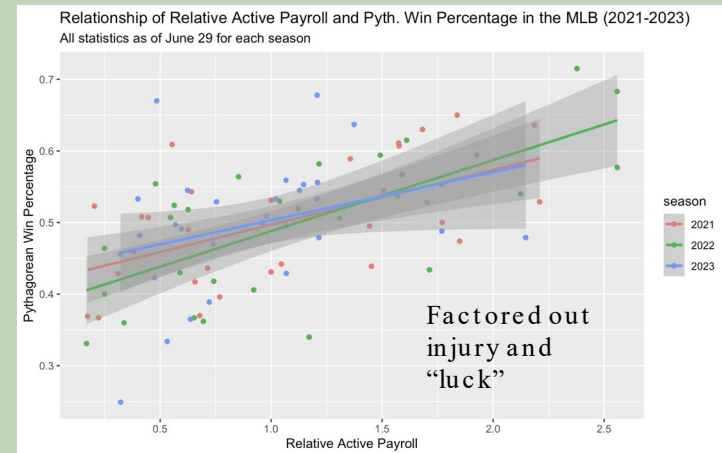
Data/ Results



Win % vs Relative Active Payroll



Pyth. Win % vs Relative Payroll



Pyth. Win % vs Relative Active Payroll

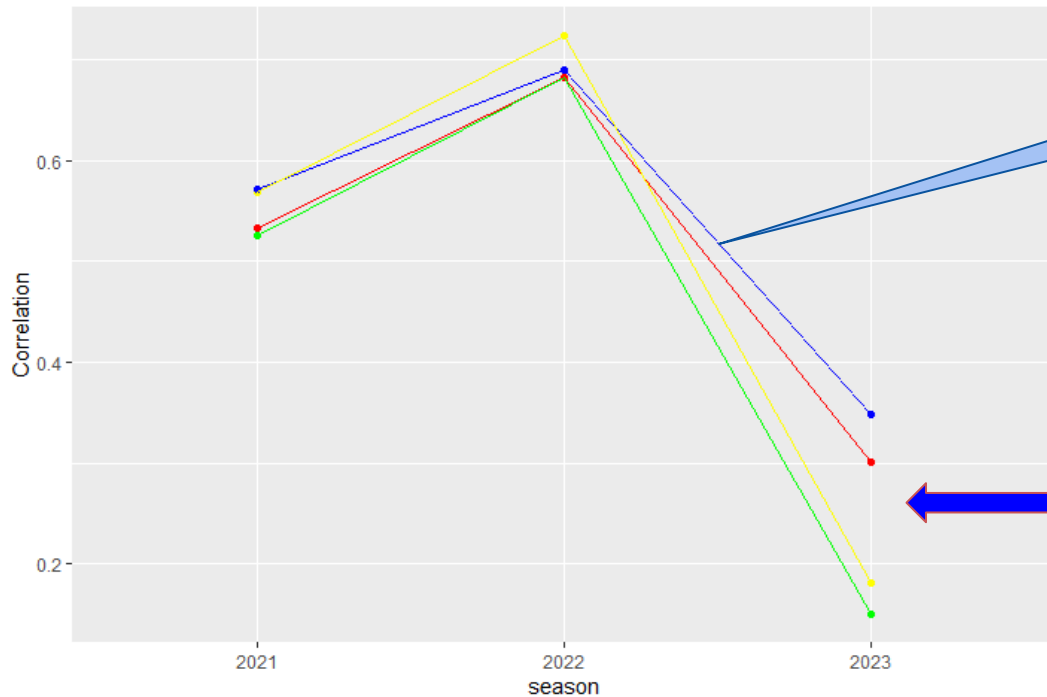
Data/Results

**Correlation Coefficient of Relationship Between
Raw/Pythagorean Win % and Yearly/Active Man
Relative Payroll**

Season	Raw vs Yearly	Pyth vs Yearly	Raw vs Active	Pyth vs Active
2021	0.5261730	0.5338243	0.5691161	0.5721851
2022	0.6829967	0.6822396	0.7233854	0.6907564
2023	0.1506257	0.3002661	0.1812102	0.3490120

Data/Results

Correlation of Raw/Pythagorean Win% and Yearly/Active Relative Payroll



Steep drop in 2023, indicating greater variability in "surprise teams"

Correlation Type

- pyth vs rel yearly
- pyth vs rel active
- raw vs rel yearly
- raw vs rel active

Significant gap caused by the "luck" factor, or the pythagorean win %

Conclusion

- The “Surprise Teams” of the 2023 MLB Season are indeed more surprising than previous years
- Correlations between win percentages and relative payrolls are significantly less in 2023 (goes from moderate to weak relationship)
- Factors such as injuries (active man payroll) and being “lucky” (pythagorean win%) do play some role. The slope of the line of best fit for 2023 becomes closer to the previous years as luck and injuries are factored out.






The Next Steps



Given more time and experience with R...

- Create a stronger, more accurate model to justify the definition of a “surprise team”
 - Take the data back longer than 3 seasons
 - Factor out as many other confounding variables as possible
 - Examine if new rule changes newly implemented for the 2023 season have also led to the greater surprise we’ve seen
 - What causes a “surprise team” to emerge?
- 
- 