"Do Big Players get the Big Results?:

The Effect of Salary Disparity on Team Performance in the MLS"

By Kabir Wagle

California Polytechnic University San Luis Obispo

Econ 464

March 17th, 2023

Professor Hamilton

I. Introduction

In this paper, I will explore the effect of wage dispersion on team performance in Major League Soccer (the leading professional soccer league in the United States).

The controversy regarding the effect of wage dispersion, or salary inequality, on team performance, stems from two competing theories: cohesion theory and tournament theory. The cohesion theory, put forward by Akerlof and Yellen (1988) states that paying everyone on a team a similar salary leads to greater team cohesion and thus better overall team performance. Sociometrically, group cohesiveness has been defined as, “within-group harmony, intermember attractiveness, or the extent to which the members reinforce each others' expectations regarding the value of maintaining the identity of the group (Stogdill 1972).” The opposing tournament theory put forward by Lazear and Rosen (1981) states that using a hierarchical wage structure encourages competition and thus betters overall team performance. The basis of this theory is that the differentiation in salary incentivizes greater efforts to obtain higher salaries, simultaneously improving team performance. These competing theories are essentially the basis of all economic studies into the matter of salary disparity.

The topic of wage dispersion within a team is something that has long been discussed by economists. It can be hard to pinpoint its exact effect on performance due to a multitude of factors including but not limited to, worker talent, management, and working conditions. However, sports teams provide an optimal environment to measure this effect. “Sports provide the ultimate avenue for examining business and management practices: owing to data availability, on the one hand, and the high degree of competition in the industry, on the other”

(Frick, Prinz, and Winkelmann 2003). On top of this, sports have a built-in measure of what is considered success, which can be uniformly measured across competitors.

In the past decades, there has been much literature published surrounding wage dispersion in both American sports leagues and European soccer leagues. In the MLB (Major League Baseball), most studies align with the cohesion theory that performance and high wage dispersion are negatively correlated (Breunig et al., 2014; Jewell & Molina, 2004; Richards & Guell, 1998). Conversely, similar studies in the NBA (National Basketball Association) such as Katayama and Nuch, 2011 found the relationship to be insignificant. Conclusions generally tend to vary on a sport-to-sport basis. This can be attributed to individual nuances within a sport. For example in Basketball, there are only 5 players on the court at a time, meaning that one individual player can have more of an impact than in say baseball, where 11 players take the field at a time.

When it comes to European soccer one of the leading studies into this matter comes from Franck and Nuesch (2011). This study examined the Bundesliga (German first division) and found a U-shaped relationship between wage dispersion and team performance. Essentially, if you found yourself with either very low or high levels of salary inequality it would improve success, however, moving toward the middle of the spectrum would begin to decrease success. Interestingly, this study supported aspects of both the cohesion theory and the tournament theory.

As you can see, professional sports are somewhat of a hotbed when it comes to researching wage dispersion. That being said, the MLS has been often overlooked given its smaller stature. Interestingly, Major League Soccer provides a uniquely ideal environment for analyzing wage dispersion. Unlike other soccer leagues around the world, the MLS has a built-in salary cap limiting the amount of money teams can spend. The one exception to this is the

Designated Player Rule, aka the Beckham Rule. This rule was introduced in 2007 as a way of allowing MLS teams to sign big-name players (such as its namesake David Beckham) in an effort to grow the stature of the league. These big-money contracts did not count toward the league's salary cap. The MLS official website states: "The Designated Player Rule allows clubs to acquire up to three players whose total compensation and acquisition costs exceed the Maximum Salary Budget Charge, with the club bearing financial responsibility for the amount of compensation above each player's Salary Budget Charge." What this has created is a unique environment where MLS teams often have 2-3 players on astronomical salaries, while the rest of the team makes significantly less. This inherent higher level of salary inequality in the MLS will make the league a very interesting sample for this study.

In this paper I will test the hypothesis that wage dispersion has no effect on team performance in the MLS, analyzing regular season data from 2007 (the year the Designated Player Rule was introduced), up until 2017.

II. Literature Review

All studies into the effect wage of dispersion on team performance in sports tend to follow a specific pattern; A) Track Data from a sport across a range of seasons. B) Measure success across those seasons. C) Measure wage dispersion across those seasons. D) Control for any external variables that may be contributing to causation. E) Model the data.

A. Tracking Data

There have been studies covering almost every major sports league in the world. Examples include; the MLB (Breunig et al., 2014; Jewell & Molina, 2004; Richards & Guell, 1998), NBA (Schouten, 2012; Simmons & Berri, 2011), NHL (Depken & Lureman, 2018), Bundesliga, (Franck and Nüesch. 2011) and Serie A (Domizio, Pellegrini, and Caruso. 2021). Almost all studies focus on a singular league, as it is difficult to compare different leagues even within a sport as they are unlikely to be economically comparable. For example, it would be inadvisable to simultaneously analyze the NBA and the CBA (Chinese Basketball Association) as they operate on different economic playing fields. The one exception to this is European soccer, as there are multiple leagues that boast similar levels of economic strength. This is seen in (Ramos 2021), where wage dispersion is measured across the top 5 soccer leagues in Europe; England, Spain, Italy, Germany, and France.

The time span of the data tended to vary across sports. For the MLB, NBA, and NHL, data was tracked for a 3-7 year period, while for any soccer league, data tended to be tracked for 8+ years. This can be attributed to the number of games played in a season. In the MLB, NBA, and NHL they play 162, 82, and 82 games respectively per season. In the top 5 European soccer leagues, they play between 34 and 38 games per season. The MLS lines up more closely with the latter, as they play 34 games a season. With this in mind, I tracked data from 2007 until 2017, an 11-year span. The data covers such a wide span that there were several new teams added to the league during this time span, altering the number of games during a season. From 2007 - 2010 there were 30 games in a season and from 2011 - 2017 there were 34 games a season.

B. Measuring Success

As mentioned earlier, sports are fertile ground for studies into wage dispersion as they have an inherent measure of what constitutes success: winning. While there are alternative methods such as Coates et al (2014), which measured success through income or profits, winning games is what sports organizations truly value (Ramos 2021).

For the MLB, NBA, and NHL, the main measures of success were regular season wins (Breunig et al., 2014; Depken & Lureman, 2018; Schouten, 2012). Depken & Lureman (2018), also incorporated goals scored and goals conceded as they wanted to examine the effects on offensive and defensive outputs. The hypothesis was that cohesion would improve defensive structure, while the individualism of high salary disparity could lead to higher attacking output.

Franck and Nüesch (2011), Domizio, Pellegrini, and Caruso (2021), and Ramos (2021) were forced to explore alternate measurements given that draws are quite common in soccer. This means you cannot look solely at wins, but instead consider overall point totals. All three of the aforementioned studies utilized some form of "percent_points", which is the number of points the team finished with that season, divided by the total points possible. I will be using a slightly adjusted version of the metric which is "points_per_game", the number of points the team finished with that season, divided by the total number of games that season. I find this metric to be more interpretable, as you can look at it as: How many points did the team earn each game? This metric can be easily transformed as percent_points = points_per_game ÷ 3.

Finally, we must consider success outside of the traditional league setting. While there aren't "playoffs" in European soccer, there are cup competitions both Domestically and Internationally that contribute to the team's definition of success for that season. Franck and Nüesch (2011), Domizio, Pellegrini, Caruso (2021), and Ramos (2021), all omitted this from their analysis. This is in part due to the lack of consistency it creates. How do you weigh percent_points in the league compared to a win in a cup competition? For that reason, I will be omitting cup competitions from my analysis as well.

C. Measuring Wage Dispersion

Franck & Nüesch (2011) and Coates et al (2014) established that the Gini coefficient is the best method for measuring salary disparity. This is the measure of income disparity that the United States Census Bureau uses and defines it as such, “The Gini coefficient ranges from 0, indicating perfect equality (where everyone receives an equal share), to 1, perfect inequality (where only one recipient or group of recipients receives all the income). The Gini coefficient is based on the difference between the Lorenz curve (the observed cumulative income distribution) and the notion of a perfectly equal income distribution” (US Census Bureau).

Additionally, Brown (1998) put forward the idea that the coefficient of variation (CV), is a good measure of income inequality. The Coefficient of Variation is simply the ratio of the standard deviation to the mean. “The CV is an attractive statistical tool since it is dimensionless” (Ramos 2021).

Lastly, studies such as Depken (2000) and Domizio, Pellegrini, and Caruso (2021) included more general measures of inequality such as the Herfindahl–Hirschman (HHI) and Theil indices. The HHI index is “a commonly accepted measure of market concentration” (US DOJ), while the Theil index “measures an entropic ‘distance’ the population is away from the ‘ideal’ egalitarian state of everyone having the same income” (US Census Bureau).

I will be considering all four of the aforementioned measures of wage dispersion: The Gini Coefficient, Coefficient of Variation, Theill Index, and HHI Index.

D. Controlling for External Variables

Wooldridge (2003) established that controls should be selected that are both useful in predicting team performance and correlated with team performance.

For obvious reasons, every study I read included some form of team salary as a control. Depken (2000) simply used total team salary, while Simmons & Berri (2010) used "relative_salary", which is the team's total payroll divided by the average payroll of all the teams in the league. The salary control was often subject to transformation, with Simmons & Berri (2010) electing for a quadratic transformation and Franck & Nüesch (2011), a logarithmic one. I will be exploring transformed versions of the "relative_salary" variable in my model.

Domizio, Pellegrini, and Caruso (2021) included a variable, "number_of_seasons", to account for the number of seasons the team has been playing in the top flight. I thought this was particularly useful for my data set given that new teams are being added to the MLS every year. There are numerous teams within the dataset where it is their very first season in the league. The other controls I will be including in my model are "year" and "number_of_players".

Ramos (2021) included a control for player value, by including the player's value as per the website Transfermarkt, and Domizio, Pellegrini, and Caruso (2021) included a control for player age. These are useful controls due to their relationship with the player's salaries. We would expect younger players to have a higher player value but lower wages, while older players to have lower player value but higher wages. These are both controls I would have liked to include however, I could not find a data source that included salaries, player value, and date of birth.

E. Modeling the Data

In every study in the realm of soccer, Franck and Nüesch (2011), Domizio, Pellegrini, and Caruso (2021), and Ramos (2021), modeled the data by looking at a measure of success, against a measure of wage dispersion, controlling for external variables.

As mentioned in the introduction, Franck and Nüesch (2011) is widely regarded as the benchmark for this topic and found “a U-formed relationship between the intra-team wage dispersion and sporting success.”

Domizio, Pellegrini, and Caruso (2021) found a “positive but inelastic relationship between relative wages and performance” and “evidence of a negative impact of pay dispersion on performance”.

Ramos (2021) aligns more closely with Franck and Nüesch (2011), putting forward that teams should either focus on an individualistic strategy (tournament theory), or a cooperative one (cohesion theory). Finding yourself in the middle of these two strategies will see a decrease in team performance.

III. Data, Methodology, and The Model

Data was measured from the 2007 to the 2017 MLS season, where each data point X_{ij} represents the i’th year of the j’th team. Overall there were 195 data points measured. Table 1 demonstrates how the number of MLS teams per season changed over time.

A. Data Tracking

Figure 1. [1,2,3,4,5,6,7,8]

No. of Teams in Major League Soccer 2007 - 2017

# Of Teams	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
22											Minnesota United
21											Atlanta United
20									Orlando City SC	Orlando City SC	Orlando City SC
19						Montreal Impact	Montreal Impact	Montreal Impact	NYCFC	NYCFC	NYCFC
18					Vancouver Whitecaps	Vancouver Whitecaps	Vancouver Whitecaps	Vancouver Whitecaps	Montreal Impact	Montreal Impact	Montreal Impact
17					Portland Timbers	Portland Timbers	Portland Timbers	Portland Timbers	Vancouver Whitecaps	Vancouver Whitecaps	Vancouver Whitecaps
16				Philadelphia Union	Philadelphia Union	Philadelphia Union	Philadelphia Union	Philadelphia Union	Portland Timbers	Portland Timbers	Portland Timbers
15			Seattle Sounders	Seattle Sounders	Seattle Sounders	Seattle Sounders	Seattle Sounders	Seattle Sounders	Philadelphia Union	Philadelphia Union	Philadelphia Union
14		San Jose Earthquakes	San Jose Earthquakes	San Jose Earthquakes	San Jose Earthquakes	San Jose Earthquakes	San Jose Earthquakes	San Jose Earthquakes	Seattle Sounders	Seattle Sounders	Seattle Sounders
13	Toronto FC	Toronto FC	Toronto FC	Toronto FC	Toronto FC	Toronto FC	Toronto FC	Toronto FC	San Jose Earthquakes	San Jose Earthquakes	San Jose Earthquakes
12	Real Salt Lake	Real Salt Lake	Real Salt Lake	Real Salt Lake	Real Salt Lake	Real Salt Lake	Real Salt Lake	Real Salt Lake	Toronto FC	Toronto FC	Toronto FC
11	NY Red Bulls	NY Red Bulls	NY Red Bulls	NY Red Bulls	NY Red Bulls	NY Red Bulls	NY Red Bulls	NY Red Bulls	Real Salt Lake	Real Salt Lake	Real Salt Lake
10	New England	New England	New England	New England	New England	New England	New England	New England	NY Red Bulls	NY Red Bulls	NY Red Bulls
9	LA Galaxy	LA Galaxy	LA Galaxy	LA Galaxy	LA Galaxy	LA Galaxy	LA Galaxy	LA Galaxy	New England	New England	New England
8	Kansas City SC	Kansas City SC	Kansas City SC	Kansas City SC	Kansas City SC	Kansas City SC	Kansas City SC	Kansas City SC	LA Galaxy	LA Galaxy	LA Galaxy
7	Houston Dynamo	Houston Dynamo	Houston Dynamo	Houston Dynamo	Houston Dynamo	Houston Dynamo	Houston Dynamo	Houston Dynamo	Kansas City SC	Kansas City SC	Kansas City SC
6	DC United	DC United	DC United	DC United	DC United	DC United	DC United	DC United	Houston Dynamo	Houston Dynamo	Houston Dynamo
5	FC Dallas	FC Dallas	FC Dallas	FC Dallas	FC Dallas	FC Dallas	FC Dallas	FC Dallas	DC United	DC United	DC United
4	Colorado Rapids	Colorado Rapids	Colorado Rapids	Colorado Rapids	Colorado Rapids	Colorado Rapids	Colorado Rapids	Colorado Rapids	FC Dallas	FC Dallas	FC Dallas
3	Columbus Crew	Columbus Crew	Columbus Crew	Columbus Crew	Columbus Crew	Columbus Crew	Columbus Crew	Columbus Crew	Colorado Rapids	Colorado Rapids	Colorado Rapids
2	Chivas USA	Chivas USA	Chivas USA	Chivas USA	Chivas USA	Chivas USA	Chivas USA	Chivas USA	Columbus Crew	Columbus Crew	Columbus Crew
1	Chicago Fire	Chicago Fire	Chicago Fire	Chicago Fire	Chicago Fire	Chicago Fire	Chicago Fire	Chicago Fire	Chicago Fire	Chicago Fire	Chicago Fire
	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017

[1] 2007: 13 teams participate in the 2007 MLS season
[2] 2008: San Jose Earthquakes join the MLS
[3] 2009: Seattle Sounders join the MLS
[4] 2010: Philadelphia Union join the MLS
[5] 2011: Vancouver Whitecaps and Portland Timbers join the MLS
[6] 2012: Montreal Impact joins the MLS
[7] 2015: Chivas USA leaves the league, Orlando City SC and NYCFC join the MLS
[8] 2017: Minnesota United and Atlanta United join the MLS

B. Response Variable: Team Performance

The MLS stores all its match data on mlssoccer.com dating back to 1996. I pulled data from 2007 to 2017 for my analysis. The dataset originally had 209 variables, most of which were either irrelevant or contained large amounts of missing data. Additionally, the data was organized by match, instead of summarized by team.

I whittled the dataset down to the variables: “year”, “home”, “away”, “home_score”, and ‘away_score”. These variables represent, the home team, away team, and respective scores for every regular season game across the 11 MLS seasons in the data set. From there I was able to calculate the overall point totals using the logic Win = 3 points, Draw = 1 point, Loss = 0 points. I used a series of ifelse() statements to assign home and away point totals, and eventually combined the two to end up with the total points for each team.

Finally, I divided the point totals by the number of games in that season; 30 games from 2007 - 2010 and 34 games from 2011 - 2017. This left me with my desired measure of team performance: points_per_game.

Table 1.

Summary Statistics of Response Variable: Points Per Game

Variable	Description	Mean	Standard Deviation	Maximum	Minimum
points_per_game	Points earned per game across an MLS season	1.37	0.291	2.03	0.471

C. Explanatory Variable: Wage Dispersion

The MLS Players Association is known for publishing player salaries in an effort to create transparency and protect players’ rights. Their website states that they mainly track two statistics: Current Annualized Base Salary and Annualized Average Guaranteed Compensation. In layman's terms, this is their base salary, as well as their salary with bonuses. The latter was solely used for my analysis as it better reflects the actual salary commitment made by the team. Data was readily available for the years 2007 - 2017.

The dataset contained the following variables: club, last name, first name, position, base salary, and guaranteed compensation. There were numerous missing values in the club sections, as well as a handful in the guaranteed compensation section. All data was recovered and manually entered using TransferMarkt[9]. Total salaries were then compiled for each team.

I used all four measures of wage dispersion on team salaries, as supported by previous literature. The Gini coefficient (gini) was calculated using the gini() function in R. The Coefficient of Variation (cv) was calculated by dividing the sd() of the team's total salary by the mean() of the team's total salary. The Theil index (thiel) was calculated using the thieil.wtd() function in R. Finally, the HHI index (hhi) was calculated using the formula seen below.

$HHI = s_1^2 + s_2^2 + s_3^2 ... + s_n^2$; where s_n = the market share of firm n expressed as an integer.

Table 2.

Summary Statistics of Explanatory Variables: Gini, CV, Thiel, and HHI

Variable	Description	Mean	Standard Deviation	Maximum	Minimum
gini	Gini Coeffcient of a team's salary a given season. Where 1 represents perfect inequality and 0 represents perfect equality.	0.533	0.147	0.873	0.311
cv	Coefficient of Variation of a team's salary a given season. Where higher scores denote greater variation between the population and the mean.	1.39	0.758	3.93	0.538
thiel	Thiel Index of a team's salary a given season. Where higher scores denote the population is further away from income equality.	0.601	0.452	0	0.144
hhi	Herfindahl-Hirschman Index of a team's salary a given season. Where 10,000 represents a Monopoly, and 0 represents Perfect Competition.	1205	945	596	5135

[9] Transfermarkt is a reputable soccer website that contains data on player information, stats, and salaries.

D. Controls

I've selected four variables to control for the effect of wage dispersion on team performance. The first is the year, which already existed in the data set. This was done to control the time series aspect of the model. The second is relative_salary, which was calculated by dividing the team's total salary by the average team salary in the league that year. This should control for the correlation between total salaries and team performance. In general, we would expect those with higher total salaries to have better team performance. Third, was the inclusion of squad_size, which will explain aspects of wage dispersion. Finally, I added a years_in_MLS variable. As seen in Table 1 many new teams have joined the league between 2007 - 2017. Including a variable for how long the team has been in the MLS will control for any effect it might have on team performance.

Table 3.

Summary Statistics of Controls: Year, Relative Salary, Squad Size, and Years in MLS

Variable	Description	Mean	Standard Deviation	Maximum	Minimum
year	The year the MLS season took place	2012	3.1	2017	2007
relative_salary	The team's relative payroll compared to the other teams in the league that year	1	0.615	3.54	0.49
squad_size	The number of players on the team that year	28.2	2.25	33	22
years_in_MLS	The number of years the team has been in the MLS	11	6.55	21	0

E. The Model

Formula 1.

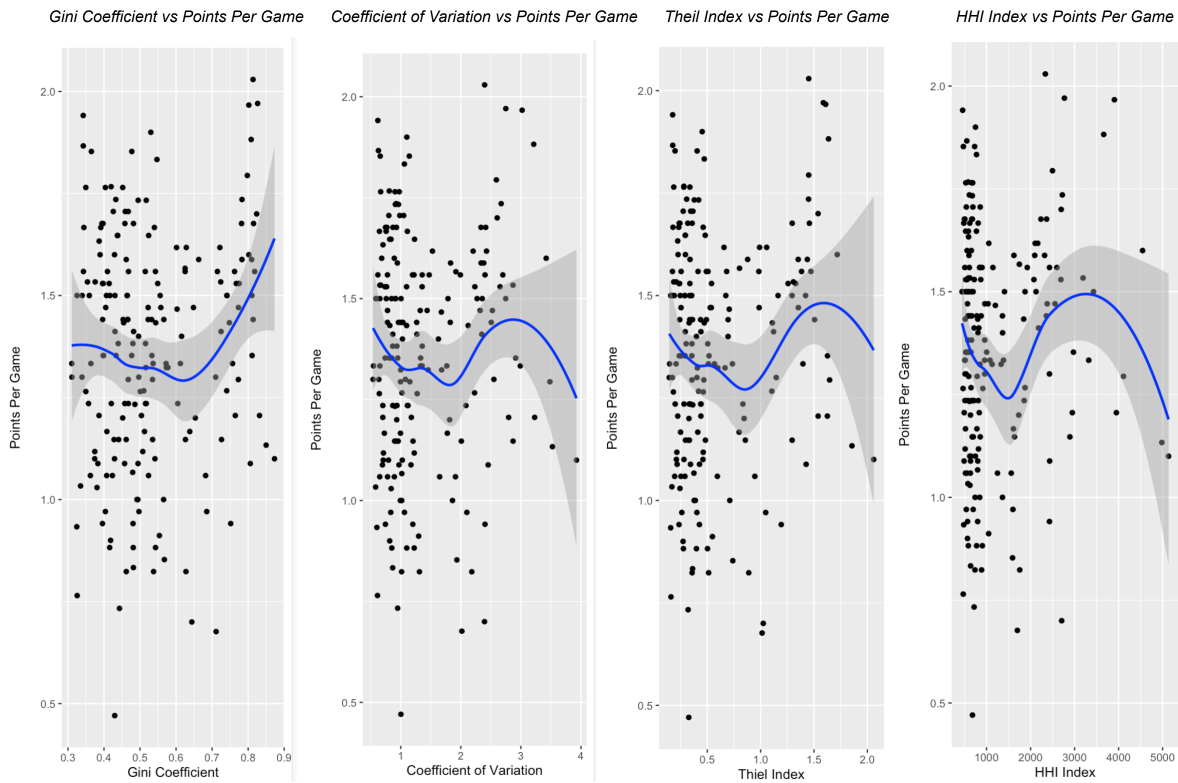
$$points\_per\_game = B_0 + B_1 dispersion + B_2 year + B_3 relative\ salary + B_4 squad\_size + B_5 years\_in\_MLS$$ [10]

Formula 1. shows the model I will be using for the analysis below. The response variable will be points per game, the explanatory variable wage dispersion, and the controls; year, relative salary, squad size, and years in MLS.

[10] dispersion variable could be any of the four measures of wage dispersion (gini, cv, thiel, and hhi)

IV. Analysis and Results

Figure 2.



Rudimentary scatterplots show no clear evidence of a linear association between the measures of dispersion (Gini, CV, Thiel, and HHI) and points per game. The data tends to be more concentrated on lower values of wage dispersion in the CV, Thiel, and HHI plots. The Gini Coefficient has a better spread of data across the x-axis. The regression line seems to be predicting some association at the tail end, however, there isn't enough data there to make a reliable prediction.

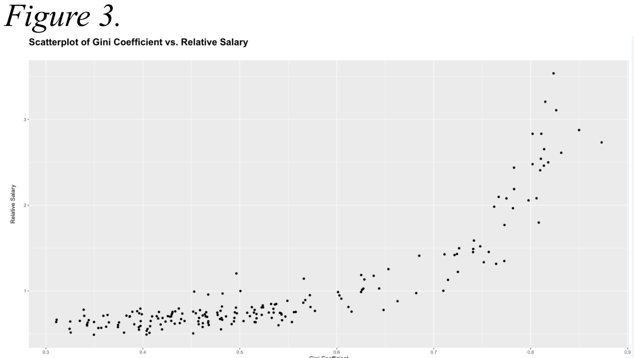
Table 4.[11]

Output of Linear Models of Wage Dispersion vs Points Per Game with Controls

Response Variable				
points_per_game				
Explanatory Variables	**Estimate**	**Std. Error**	**t statistic**	**P-value**
gini	-0.489	0.270	-1.815	0.071*
cv	-0.114	0.048	-2.397	0.0175**
thiel	-0.223	0.103	-2.169	0.0314**
hhi	-0.07797	0.038	-2.065	0.0402**
Controls				
year				
relative_salary				
squad_size				
years_in_MLS				
Significance Codes: 0.01*, 0.05**, 0.10***				

Table 4. shows the output of the full linear model for each individual measure of wage dispersion. All four measures have a significant negative association with an increase in points per game. Gini at the 10% significance level and CV, Thiel, and HHI at the 5% significance level. As hypothesized, the relative_salary variable was highly significant (***) in each of the four models. Across all the models, the year and squad_size controls were highly insignificant and are at risk of being dropped from the model.

Figure 3.



As outlined in Domizio, Pellegrini, and Caruso (2021), given the inherent logarithmic relationship between wage dispersion and relative salary, it makes sense to include a logarithmic transformation of the relative_salary variable.

[11] The HHI variable was divided by a factor of 1000 for the model

Table 5.[12]

Output of Transformed Linear Models of Wage Dispersion vs Points Per Game with Controls

Response Variable				
points_per_game				
Explanatory Variables	**Estimate**	**Std. Error**	**t statistic**	**P-value**
gini	-0.654	0.325	-2.012	0.0457**
cv	-0.142	0.055	-2.581	0.0106**
thiel	-0.264	0.122	-2.159	0.0321**
hhi	-0.080	0.041	-1.966	0.0507*
Controls				
log(relative_salary)				
years_in_MLS				
Significance Codes: 0.01*, 0.05**, 0.10***				

Including the logarithmic transformation of relative_salary and dropping the year and squad_size controls led to changes in the significance of the wage dispersion variables. Gini had the greatest increase in significance, now falling within the 5% significance level. The significance of CV slightly increased, while the significance of Thiel slightly decreased. HHI actually moved outside the 5% significance level.

Although the transformation didn't ubiquitously improve significance, I am particularly interested in its effect on Gini and CV. Gini as it is the most widely accepted measure of wage dispersion, and CV as it is the most significant measure in the model.

The slope coefficient associated with the gini variable indicates a negative linear association between gini and points_per_game. Meaning, for teams in the MLS between 2007-2017, as the level of wage dispersion, or salary disparity, within a team increases, there is an estimated negative effect on overall team performance, after controlling for relative salary and the number of years the team has played the MLS.[13]

That being said, if you wanted to build a truly predictive model, the coefficient of variation would be the best indicator of effect on team performance for this data set.

[12] See Table 6. in the appendix for the full output.
[13] Formula 2. was used to check the assumptions of linear regression as seen in Figure 4.

V. Conclusion

Economists and statisticians have long researched the effects of wage dispersion, or salary disparity, on performance, through the lens of professional sports. In this study, I aimed to explore this relationship in Major League Soccer, the leading professional soccer league in the United States.

The logarithmic relationship established in Franck and Nüesch (2011), Domizio, Pellegrini, and Caruso (2021), and Ramos (2021) was found to be insignificant in this dataset in comparison to the linear relationship seen in Tables 5 and 6. This provides strong evidence in support of the cohesion theory. MLS teams that paid their players a similar salary between 2007 - 2017 would have expected to see an increase in their points per game across the season.

Why did results in the MLS differ from those in the top European Leagues (Premier League, Serie A, La Liga, Bundesliga, Ligue 1)? I would assume this had to do with the previously mentioned Designated Player Rule, which caused almost every MLS team to have relatively high levels of salary disparity compared to the rest of the world. I believe this allowed me to reach a stronger conclusion than previous studies.

What does this mean for the MLS? General Managers, Technical Directors, and Coaches should really consider the effects of choosing to break the bank for big-name players over having a balanced squad. Currently, it is the beginning of the 2023 season, and since 2017 there has been a concerted effort by MLS clubs to build more balanced squads. Rather ironically, in the 2022 MLS Cup Final, a transfer-savvy, well-balanced, Philadelphia Union team, lost to the superstar-spearheaded LAFC. Perhaps big-name players can get those big results, however, big money definitely does not equal big results.

There is further investigation to be done into the topic of wage dispersion and team performance in Major League Soccer. Most glaringly, a newer sample size including up to the 2022 season. Additionally, there could be more holistic measures of team performance explored through cup competitions such as the US Open Cup and CONCACAF Champions League. Finally, a broader range of control variables could be added to the model such as player value, player age, and player experience.

In the context of Economics, it is difficult to say how much these findings can be extrapolated outside of soccer. However, you can make some generalizations to do with cohesion theory. Namely, bringing talented people into a team on high wages may not improve team performance, if it is at the expense of the overall squad harmony. Any sort of management team looking to maximize efficiency would do well to keep that information in mind.

Appendix

Table 6.

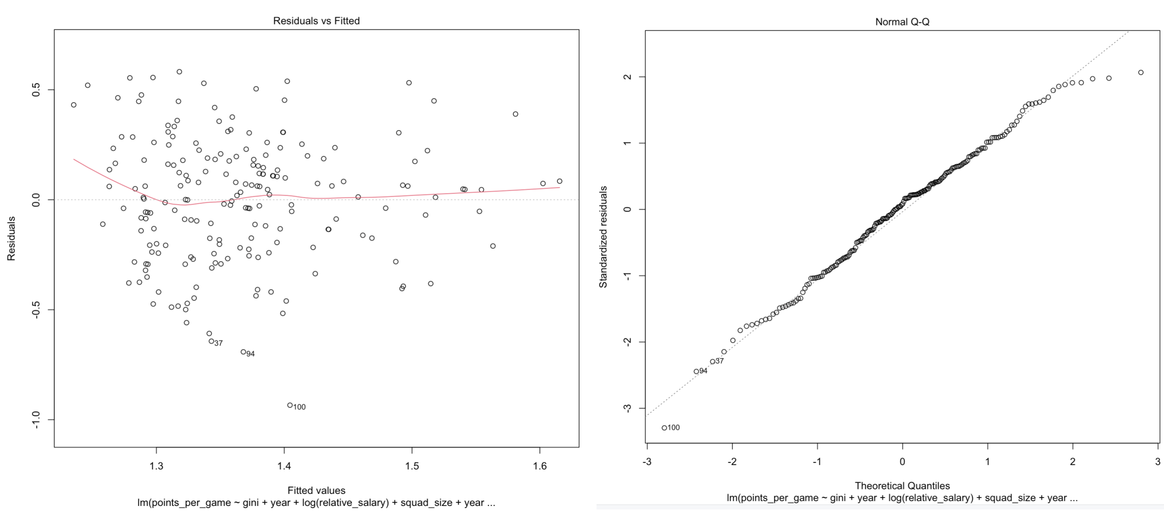
Full Ouput Transformed Linear Models of Wage Dispersion vs Points Per Game with Controls

Response Variable						
points_per_game						
			Test Statistic (P-Value)			
Explanatory Variables			**1**	**2**	**3**	**4**
gini			-2.012 (0.0457**)			
	Estimate	-0.654				
	Std. Error	0.325				
cv				-2.581 (0.0106**)		
	Estimate	-0.142				
	Std. Error	0.055				
thiel					-2.159 (0.0321**)	
	Estimate	-0.264				
	Std. Error	0.122				
hhi						-1.966 (0.0507*)
	Estimate	-0.080				
	Std. Error	0.041				
Controls						
log(relative_salary)			2.946 (0.0036***)	3.547 (0.00049***)	2.980 (0.00326***)	3.065 (0.00249***)
years_in_MLS			1.586 (0.11448)	1.326 (0.186437)	1.544 (0.12414)	1.360 (0.17557)
Significance Codes: 0.01*, 0.05**, 0.10***						

Formula 2.

$$points\_per\_game \ = B_0 \ + \ B_1 gini \ + \ B_2 log(relative\_wages) \ + B_3 (years\_in\_MLS)$$

Figure 4.



References

Akerlof, G.A., Yellen, J.L., 1988. Contemporary Economic Policy, 40(3), 513–525. “The Fair Wage-Effort Hypothesis and Unemployment.” The Quarterly Journal of Economics 105, no. 2 (1990): 255–83.

Berri, D. J. and Jewell, T. R. (2004) Wage inequality and firm performance: professional basketball’s natural experiment, Atlantic Economic Journal, 32, 130–9

Breunig, R., Garrett-Rumba, B., Jardin, M., & Rocaboy, Y. (2014) Wage dispersion and team performance: a theoretical model and evidence from baseball, Applied Economics, 46:3, 271-281.

Brown, C. E. (1998). Applied multivariate statistics in geohydrology and related sciences. Berlin: Springer.

Bureau, U. S. C. (2021, October 8). *Gini Index*. Census.gov, from https://www.census.gov/topics/income-poverty/income-inequality/about/metrics/gini-index.html

Coates, D., Frick, B., & Jewell, T. (2016). Superstar Salaries and Soccer Success: The Impact of Designated Players in Major League Soccer. Journal of Sports Economics, 17(7), 716–735.

Di Domizio, M., Bellavite Pellegrini, C., & Caruso, R. (2022). Payroll dispersion and performance in soccer: A seasonal perspective analysis for Italian Serie A (2007–2021).

Franck, E. P., & Nuesch, S. (2011). The effect of wage dispersion on team outcome and the way team outcome is produced. Applied Economics, 43, 3037-3049.

Frick, B., Prinz, J. & Winkelman, K. (2003) Pay inequalities and team performance: empirical evidence from the North American Major Leagues. International Journal of Manpower, 24(4), 472–488.

Gasparetto, T., & Barajas, A. (2022). Wage Dispersion and Team Performance: The Moderation Role of Club Size. Journal of Sports Economics, 23(5), 548–566.

Herfindahl-Hirschman Index. The United States Department of Justice. (2018, July 31), from https://www.justice.gov/atr/herfindahl-hirschman-index

Katayama, H. & Nuch, H. (2011) A game‐level analysis of salary dispersion and team performance in the national basketball association. Applied Economics, 43(10), 1193–1207.

Lazear, E. P. (1989) Pay equality and industrial politics, Journal of Political Economy, 97, 561–580.

Mlssoccer. (n.d.). Roster rules and regulations. mlssoccer, from https://www.mlssoccer.com/about/roster-rules-and-regulations

Ramos, M. M. F. B. (2021). The Impact of Wage Dispersion in Football Teams Performance: An Overview of the Big-5 European Leagues of Football. NIMS - Data Science and Advanced Analytics.

Richards, D.G. & Guell, R.C. (1998) Baseball success and the structure of salaries. Applied Economics Letters, 5(5), 291–296.

Simmons, R., & Berri, D. J. (2011) Mixing the princes and the paupers: Pay and performance in the National Basketball Association, Labour Economics, Volume 18, Issue 3, 381-388

Stogdill, R.M. (1972) Group productivity, drive, and cohesiveness. Organizational Behavior and Human Performance, 8(1), 26–43.

Wooldridge, J. M. (2003) Introductory Econometrics: A Modern Approach, 2nd edn, South-Western, Cincinnati, OH.