Drew University

College of Liberal Arts

What is the Relationship between Pay Dispersion and Team Productivity in the National Basketball Association?

A Thesis in Economics

by

Lisa Cappiello

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Bachelor in Arts

With Specialized Honors in Economics

May 2020

Dedication: This study is dedicated to my parents Catherine Fama and Sal Cappiello. I would not be the person I am without you two. I would also like to dedicate this to my amazing committee (Maliha Safri, Jennifer Kohn, and Chris Andrews) for sticking with me through this long but rewarding journey.

Abstract

This honors thesis seeks to investigate the relationship between pay dispersion and team productivity within the National Basketball Association. I take data from a variety of sources to create an excel spreadsheet, where I calculate the Gini coefficient. I decide that team conference and year are important independent variables for this study. I use four regression techniques: ordinary least squares (OLS), lagged dependent variable, fixed effects and random effects. I use these four models because the basic OLS has potential violations of assumptions which the other regressions relax. I ultimately no statistically significant relationship between pay disparity and team productivity within the National Basketball Association.

Table of Contents

Introduction

How much does team cohesion affect productivity and success? I examine this question in the context of the National Basketball Association (NBA) looking at how team cohesion works in the presence of great income inequality. The NBA is a natural place to examine this question because there is only one ball and five players on the court; therefore winning requires a great amount of teamwork. On the other hand, players may have incentives to play selfishly in order to increase their personal performance and thereby increase their salary. In this paper, I examine the association between wage disparity and team success through econometric methods and find that there is no statistically significant relationship between wage disparity and team productivity.

Specifically, I hypothesize that if a team has larger wage disparity, they will not win many games and consequently not make playoffs or win a championship. As an unpaid college basketball player, I am interested in basketball and as an economics major I am interested in the determinants of productivity, and in particular, I am interested in how it is that team comradery is both weakened and strengthened. I hope to understand what affects a team's capacity to achieve cohesion, since “team owners and managers should be aware of the potential tradeoff between winning and the superstar effect” (Annala and Winfree 2011). The superstar effect describes the idea that a talented “superstar” player comes to an organization, and the game becomes all about them. They play selfishly, and consequently the team does not win many games. For example, Russell Westbrook of the Oklahoma City Thunder has been a consistent star in the league but has not won a single championship. He is also consistently paid a great deal from the Oklahoma City Thunder franchise compared to lower paid players. Seeing the wide difference in wages for different players on the same team led me to wonder if that had any relevance to achieving the overall goal of winning. For professional athletes unlike amateur athletes, it is not only about the game anymore, this is how they make their living and support their families.

Sports are a good context to examine the relationship between productivity and team cohesion for three reasons. First, productivity in sports can be put in black and white terms: winning determines productivity and losing indicates its absence. Second, players face a clear tradeoff between personal performance and team performance. Basketball has limits within the game, much like any other sports. There is only 40 minutes, one ball, and a finite amount of players on the court at once. There is an element of codependence on teammates in order to produce in basketball. According to Simmons and Berri (2011) "In basketball, shot attempts in a game are finite; if a player takes more shots, his teammates must take fewer shots since there is only one ball and limited playing time." These limits are of importance because there are factors that include opportunities for a player to perform selfishly. Lastly, the data is publically and readily available for use.

In order to test this hypothesis, I use data from HoopsHype, a well-known basketball website that lists official statistics and salaries for all 30 NBA teams for the years 2009-2018. I operationalize the variable of wage inequality using the Gini coefficient for the salaries published by HoopsHype. I chose the Gini coefficient because it is widely used to understand inequality in many contexts. For example, the U.S. Census Bureau uses it to depict income inequality as do numerous well-cited publications in sports economics such as Coates, Frick and Jewell's (year) "Superstar Salaries and Soccer Success: The Impact of Designated Players in Major League Soccer".

I also use data from LandofBasketball, another basketball super site with comprehensive NBA team records, as well as other relevant statistics. These records will be used for the same years 2009-2018. I took the data on the number of wins for the entire season - including if they played any playoff games or won the championship - and what conference a team was competing in that season.

The empirical analysis will begin with an ordinary least squares (OLS) model, with the dependent variable of wins and the following independent variables: Gini coefficient to depict wage disparity, the team's conference which is defined by the EastWest variable, and year. I use the basic OLS, OLS with a lag, fixed effects and random effects models. It is important to use these four models to handle different underlying assumptions that may be violated by the basic OLS model. I evaluate the OLS assumptions with residual graphs. The methodology section will detail assumptions underlying each model, and the results section will provide the results found from each model's regression including diagnostic graphs for these assumptions.

This study is examines the most recent ten seasons for all 30 teams in the NBA. While previous works have looked at other sports like baseball (Annala and Winfree 2010) and hockey (Gomez 2002) in the search of pay disparity effects on productivity, this study differentiates itself by using very recent NBA data as well as looking at the productivity of a team as opposed to the productivity of a player. Other studies that have been done on the NBA, Berri and Jewell (2004) and Katayama and Nuche (2011), also did not find any significant impact of pay dispersion in their studies. To differentiate my work from previous studies, I use team wins instead of winning percentage, study a different time period (2009-2018) and use an econometric methodology.

My work draws heavily on the literature emphasizing cohesion which would predict negative effects of wage inequality on team productivity, rather than tournament theory which would predict positive effects of wage inequality on team productivity. Ultimately however, I find that wage inequality within a team has insignificant effects on a team's wins. In none of my models did the Gini coefficient produce a statistically significant impact.

This thesis is divided into six sections. The first section is the literature review, where I describe previous works on Gini coefficient, Productivity, Cohesion and Tournament Theory. The second section describes the data, each source and what each variable measures. The third, explains the summary statistics of the data set that can help illustrate the wage dispersion and winning environment of the NBA. The fourth section is the method section, where I explain the six basic OLS assumptions and how I will control for unquantifiable events of sports. The fifth section, will explain the results from these regressions. Lastly, the conclusion will leave you with the main points of this paper as well as what other avenues are for future research.

Literature Review: Does Inequality Matter for Productivity?

In this section, I describe the most important papers that I refer to in my study. When doing so, I explain the main purpose of each paper, the relevant results, and how those results influenced my own methods, data, and understanding of the subject at hand. The section begins by explaining works on the Gini coefficient, then defining what productivity is in sports: wins. Lastly, I explain the two recurring theories that have appeared in other case studies examining productivity and wage dispersion within sports teams: first, *cohesion theory* and second, *tournament theory*.

Using the Gini Coefficient to Measure Inequality

Many studies have used the Gini coefficient to measure inequality in many different settings including, the U.S. Census Bureau's studies on U.S. income inequality, Coates Frick and Jewell's study on Major League Soccer (2014), and Annala and Winfree's study of Major League Baseball (2011). Those two case studies which I mirror in this study clearly show that Gini coefficient is one way to capture inequality through quantitative means.

There are two ways to go about calculating the Gini itself: parametric and non-parametric methods. Slottje (1989) examines the difference between the two methods and concludes, "sometimes the parametric method dominates and sometimes the non-parametric method dominates" (p. 196). The parametric method assumes the pay distribution for the team is normally distributed, whereas the non-parametric method does not assume the pay follows any prescribed distribution. By not assuming the pay distribution to be normal, the non-parametric method allows the data to determine the distribution itself.

In this study I use the parametric method to calculate the Gini because Slottje's (1989) paper concludes with a quote about a number of noted econometricians stating, "... if forced to do it one way, then they would do it parametrically, even though empirical evidence is split." (p. 196). Following this method as well as graphing my own data where the histograms show normal distribution, I will be using the parametric method.

Productivity

Productivity in sports can be easily defined as a team's number of wins. The only possible outcomes are winning and losing, there is no in between. Especially in basketball, there are no ties between opponents: there can emerge only one winner and one loser in each contest. In Annala and Winfree's (2011) study "Salary distribution and team performance in Major League Baseball", 2011, productivity is defined by winning percentage. In Coates and Fricks' (2016) study, they explain productivity as league points per game, but in soccer their points are not as easy to come by and the league structure is a bit different than that of the NBA.. In this study, I choose the number of wins as opposed to winning percentage because I feel it is important to note what teams play more by making it to the playoffs and get to play in the championship. By using winning percentage it allowed all teams to be held at a more standard level, which allows for teams who make the playoffs to be held at the same standard as those who did not make the playoffs. By holding these teams to the same standard we lose the importance of making it to a postseason and playing more games. In my study, I include regular season, playoffs and championship games. If a team is producing they will play more games, and for that reason I felt it was important to include playoffs and championships in my wins variable because by making it to the championship that team has surpassed the rest in amount of games played.

Cohesion Theory

Cohesion theory is defined by Levine (1991), "Firms want to narrow wage dispersion in order to increase group cohesiveness and productivity" (p. 1). Cohesiveness in this theory suggests more equal pay across workers improves productivity. *Cohesion* theory emphasizes horizontal dispersion, which is a worker to worker dispersion (workers at the same level in an organization), rather than the relation between CEO (Chief Executive Officer) to worker. Levine (1991) explains the importance of interdependence of productivity in *cohesion theory.* Interdependence appears because if you are under an organizational strategy that allows a worker to be compensated for their own work, they are less inclined to work with others. But if the worker's productivity is based on group rather than individual performance, they are more inclined to work with one another in order to achieve the goals of the firm. Under *cohesion theory*, it is important for workers to feel as though they are the same. When groups rely on group norms to ensure high levels of effort, a more cohesive pay structure plays a role (Levine, 1991). A feeling of being a part of something larger is also key in this theory, "the perception of social equity and the perception that everyone is a full member of the enterprise" (Lawler, 1981, p. 225).

Cohesion theory is exemplified through a study done by Annala and Winfree (2010) on Major League Baseball. In their study, they used panel data from 1985 to 2014. They find "evidence that payroll inequality within a team is negatively related to on field performance, in terms of regular season team winning percentages in Major League Baseball." In other words, their paper suggests that the MLB is characterized by *cohesion theory*, (Annala Winfree, 2010).

The MLB and NBA are similar in that they are team, not individual sports. Although, in the MLB there is only one batter at a time, and the lineup rarely changes. However, each player must rely on the other to bring them home to score runs. Similarly, in basketball you need your teammates to do well in order to win the game, there is only one ball and a finite amount of shots in a game. Since one person cannot make all shots, there has to be trust in other teammates to do so as well.

One case where cohesion theory may speak to basketball is the case of the 2012 San Antonio Spurs. The San Antonio Spurs had a relatively more horizontal pay dispersion with a Gini coefficient of .292,. In 2012 the Spurs led the league in assists per game with 25 assists. This statistic alone speaks to the strategy of the Spurs, to always make the next pass to the open player, maintaining team interdependence. That year they had won 73 games and lost 30, making it to the NBA Finals and losing to the Miami Heat. This great record, along with a very equal Gini coefficient makes the San Antonio Spurs a potential example of *cohesion theory*.

Tournament Theory

The other theory - *tournament theory* - is defined by Shaw (2014) as "the prediction that higher pay dispersion will be associated with higher performance … High pay dispersion will on balance increase effort levels, as steep increases in prizes across organizational levels will facilitate greater effort and competition" (p. 523)". Shaw (2014) explains what tournament theory is and in what ways has potential to work. There are four components that go along with *tournament theory*. The first component is vertical wage dispersion. *Tournament theory* is most commonly seen in vertical dispersion studies. Vertical dispersion is most commonly noted as CEO to worker dispersion, or workers at different levels within the same organization. The second is sorting. When firms follow a *tournament theory* strategy there is a tendency for workers to sort into two groups. The first group includes the average or above average workers who will become motivated to raise their status in the ranks of the firm. The second group includes those who are average or below and they become uninterested in being at the firm, look for jobs elsewhere. The third component is dispersion “driven by legitimate dispersion-creating practices”. This is intentional in hopes of these practices being effective motivators. In sports, these practices can be used as a motivating factor to lower paid players. But for players who feel as though they have proved themselves and are offered a lower contract than they feel they deserve, they may be less inclined to go to that team or work hard for that team. But from my experience, statistics do not tell you everything about a player, there is much more or much less a player can bring to a team other than their statistics. For example, Shane Battier was known as the “No-Stat All-Star”. Battiers statistics were average for the NBA, but he proved his impact every game by guarding the other teams best player. He may not have been the biggest offensive threat, but he made a difference in the game that did not show up in the box score. The last component is, how tournament theory relates to the “superstar effect.” The superstar effect is when a team brings in a high profile player, and instead of the team’s winning, the game becomes all about the one superstar player. Russell Westbrook at Oklahoma City is an example. Westbrook is a great player, and although he does well and brings the Thunder franchise revenue, he has not won a single championship with the franchise. Westbrook takes the majority of the shots and opposing teams know to stop OKC they only need to stop him.

Ding et al (2009) and Shaw (2002) both did studies on factory organizations. Each study defined vertical dispersion. When looking at vertical dispersion Shaw (2002) and Ding et al (2009) found positive effects, workers maintained safe conditions, and there was greater production. In more plain words, *tournament theory* predicts that wage inequality can impact productivity in positive ways, in contrast to *cohesion theory*.

Simmons and Berri (2010) use data from 1990 to 2008 from the NBA and find *tournament theory* describes accurately their results. They study pay dispersion in relation to players' individual statistics and team productivity. This intrigued me because it also was done on the NBA but looked at individual player success rather than team wins. The results of their study, "inequality appears to have a positive impact on player performance in the NBA" (p. 388). They suggest players individual productivity can be improved because of inequality, but my focus is different and on team productivity.

One team whose experience may be characterized by tournament theory is the 2018 Houston Rockets, which had a high level of pay dispersion but made it all the way to the NBA playoffs behind star James Harden. Their Gini coefficient was .743 showing very large wage gaps within their pay structure. The Rockets record that year was 53 wins and 29 losses, while taking the two time defending champions, Golden State Warriors to game 7 in the playoffs. Harden carried the team that year through hardship i.e. early season injuries and mid season suspensions. He was a prime example of a superstar coming through and carrying his team to success.

Data

The data I collect for this study is taken from a variety of sources. For each variable I explain the code and its source. The coding and sources of each variable can be found in Table 1.

Table 1

Teams			Years	
0	Atlanta Hawks		0	2009
1	Boston Celtics		1	2010
2	Brooklyn Nets		2	2011
3	Charlotte Hornets		3	2012
4	Chicago Bulls		4	2013
5	Cleveland Cavaliers		5	2014
6	Dallas Mavericks		6	2015
7	Denver Nuggets		7	2016
8	Detroit Pistons		8	2017
9	Golden State Warriors		9	2018
10	Houston Rockets			
11	Indiana Pacers		**East/West**	
12	LA Clippers		East	0
13	LA Lakers		West	1
14	Memphis Grizzlies			
15	Miami Heat		**Playoff**	
16	Milwaukee Bucks		No	0
17	Minnesota Timberwolves		Yes	1
18	New Orlens Pelicans			
19	New York Knicks			
20	Oklahoma City Thunder			
21	Orlando Magic			
22	Philadelphia 76ers			
23	Pheonix Sun			
24	Portland Trail Blazers			
25	Sacramento Kings			
26	San Antonio Spurs			
27	Toronto Raptors			
28	Utah Jazz			
29	Washington Wizards			

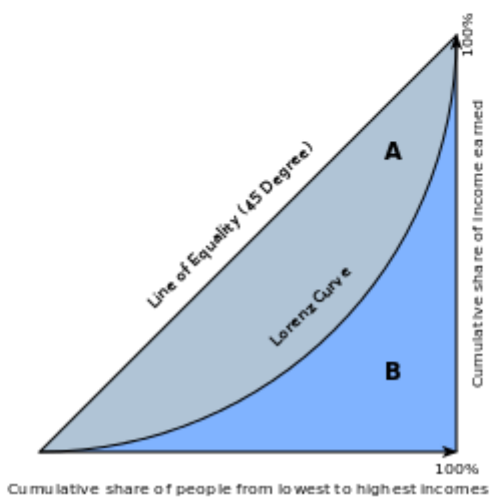
First, I take individual salary data for each player to calculate the Gini coefficient. I use HoopsHype USA Today Sports, a basketball website, as the source to obtain all players salaries. The independent variable, the Gini coefficient is a calculation based on each team's salary of each rostered player for 2009-2018. Ten years of player salaries (about 15-20 players on each team every year) for each team, 30 teams, which came out to 5,419 observations in total.

In this study the salaries are based on the end of season rosters, all trades included. These salaries are base salaries and do not include signing bonuses given to the players. Bonuses are not published widely, with only a few cases coming to light in the press . For example, Stephen Curry signed a $201 million dollar- 5 year contract, but is getting $34 million dollars each year, therefore salary data cannot capture his bonus. Rosters included players who were in “two-way contracts” between the NBA and G-League, which is the minor league equivalent for the NBA. Two-way contracts are for players who are on both the NBA and G-League rosters; these players spend most of their time in the G-League but are also allowed to spend 45 days with their NBA team. Players have options on their contracts of continuing on, renegotiating, getting extensions on their current contracts, getting another standard contract, or getting traded. No player is counted twice and in- season trades cannot occur after February. Each team makes about 1 or less in season trades per year. There are special exceptions to this, but these rarely occur i.e. this year's (2020) Golden State Warriors. Although it is also important to note bonus data was not at all available, it was impossible to include it in my examination of inequality, even though it would worsen inequality measures. I also did not use current U.S. dollars adjusted for inflation salaries because I am just looking for the dispersion, so using nominal or real salaries would not have any effect when calculating the Gini coefficient. It is also important to note that the NBA follows a soft salary cap structure where owners/front-office executives can choose to go over their salary caps and pay a luxury tax on the additional money used. They can offer contracts that are greater than their salary cap to obtain a specific player they are targeting if they feel he is worth the luxury tax. By collecting the salary data in this fashion, including two-way players, and not including bonus or incentive money, I acknowledge my data was limited by what is publicly available, as well as has potential for scale issues due to the inclusion of two-way players.

After calculating the Gini coefficient I have 300 observations from 30 teams over 10 seasons. With these components, player name, team name, salary, and year I was able to calculate the Gini coefficient for each team for each respective year by using STATA's function ginidesc.

$$□□□□ = □ / (□ + □)$$

The Gini coefficient basically measures how far an actual pay dispersion differs from a perfectly equal pay dispersion. The Lorenz curve is a graphical representation of the cumulative distribution function. Where A is the area above the Lorenz curve but under the Line of Equality and B is the area underneath the Lorenz Curve. The Gini coefficient is the percentage (A) of the total area (A+B).



In our case the percentage of cumulative players for each team/year is on the X-axis and percentage of cumulative income is on the Y-axis. A Gini coefficient of 0 explains a completely equal structure, while a Gini coefficient of 1 indicates a completely unequal dispersion.

The dependent variable Wins, was taken from LandofBasketball.com, a basketball site with the history of all NBA teams records.. A team's ultimate goal is to win more games than they lose. To be productive and achieve the goal, they must win games while forcing opponents to lose. The Wins variable includes regular season wins as well as playoff and championship wins. Every team plays 82 games in a regular season. The maximum number of games one team can play is 110, including 82 regular season and 28 postseason (taking every series to 7 games). The minimum amount of games a team can play is 82. As for playoff structure the best 8 teams of the Eastern Conference and the top 8 teams in the Western conference get to play in the postseason. The 8 teams from the East play one another until they are down to one winner, and the same goes for the West. The winner from the Eastern Conference then plays the winner of the Western Conference. Below is a sample of a standard NBA playoff bracket.

"NBA Playoffs Bracket 2019: Full Schedule, Dates, Times, TV Channels For Conference Finals". *Sportingnews.Com*, 2020, https://www.sportingnews.com/us/nba/news/nba-playoffs-bracket-2019-schedule-dates-times-tv-channels-conference-finals/1sjww7k5mndy2158vb85n2eot3.

The binary variable I included was EastWest. The EastWest information was taken from the NBA's official website. It indicates what conference the team based on geographic locations. This variable I believe potentially has an effect on productivity of winning a game and ultimately a championship because when placing people into playoffs it is based on the top teams in the

East conference versus those top teams in the West conference. A team can qualify for playoffs by being in the top 8 of their respective conference, having above a .500 record, and beating teams with similar records who could also qualify for a playoff seat (if teams are tied in record they look at their regular season matchups). This could be of importance in my study because, if the West conference is far more competitive and talented, the teams that did not make the West conference playoffs could potentially beat those top teams who made the playoffs for the East conference. For example, in 2016 the Portland Trailblazers of the Western Conference were ranked higher overall for the league and did not make the playoffs with a 41-41 record, but the Milwaukee Bucks of the Eastern Conference were ranked below them and did make playoffs with a 42-40 record. Throughout the regular season each team will play 30 games in their own respective conference and 52 non conference games. The reason I include playoff wins is because a team's objective is to obtain as many wins as possible which include playoffs and a championship.

There was one unusual season from 2011-2012, when there was a stalemate between the league and the players association. This lockout was the fourth lockout in NBA history and was the result of the 2005 CBA (Collective Bargaining Agreement) expiration. This lockout lasted 161 days and delayed the start of the 2011-2012 season. This could have potential effects on my study because each team that year did not play the full 82 game regular season schedule, instead playing 66 games. I ran regressions both including and excluding the 2011-2012 season, and found there is no statistical difference between the two results. Therefore, for the above reason, I keep the 2011 season in my data set.

Summary Statistics

Tables 2 through 7 depict the summary statistics. Tables 2 through 4 depict the summary statistics of the independent variables, while Tables 5 through 7 depict the summary statistics of the dependent variable. These statistics are for all 30 teams for all 10 years, 2009-2018.

Gini Coefficient	Dispersion Statement
$G < 0.2$	Income Equality
$0.2 < G < 0.3$	Relative Equality
$0.3 < G < 0.4$	Adequate Equality
$0.4 < G < 0.5$	Big Income Gap
$0.5 < G$	Severe Inequality

I define the Gini coefficient through year, team, and conference to determine if there were any significant dates or teams skewing the results. For the independent variable Gini coefficient, the average Gini for all the teams and all years is .546, showing that pay dispersion for NBA teams tends to be severely unequal.

Table 2

Team	Observations	Mean	Std. Dev.	Min	Max
Atlanta Hawks	10	0.558	0.052	0.501	0.637
Boston Celtics	10	0.562	0.077	0.463	0.679
Brooklyn Nets	10	0.569	0.087	0.378	0.656
Charlotte Hornets	10	0.589	0.047	0.521	0.656
Chicago Bulls	10	0.469	0.061	0.39	0.573
Cleveland Cavaliers	10	0.552	0.071	0.458	0.669
Dallas Mavericks	10	0.57	0.088	0.429	0.732
Denver Nuggets	10	0.519	0.075	0.406	0.664
Detroit Pistons	10	0.495	0.062	0.418	0.624
Golden State Warriors	10	0.559	0.071	0.49	0.691
Houston Rockets	10	0.588	0.106	0.451	0.743
Indiana Pacers	10	0.515	0.056	0.438	0.602
LA Clippers	10	0.549	0.059	0.464	0.65
LA Lakers	10	0.611	0.073	0.456	0.7
Memphis Grizzlies	10	0.599	0.024	0.565	0.644
Miami Heat	10	0.515	0.072	0.377	0.617
Milwaukee Bucks	10	0.619	0.059	0.528	0.729
Minnesota Timberwolves	10	0.46	0.109	0.215	0.612
New Orlens Pelicans	10	0.604	0.054	0.512	0.676
New York Knicks	10	0.607	0.052	0.505	0.685
Oklahoma City Thunder	10	0.547	0.099	0.387	0.697
Orlando Magic	10	0.51	0.059	0.38	0.565
Philadelphia 76ers	10	0.559	0.059	0.443	0.629
Pheonix Sun	10	0.53	0.072	0.378	0.638
Portland Trail Blazers	10	0.502	0.051	0.421	0.614
Sacramento Kings	10	0.478	0.091	0.294	0.601
San Antonio Spurs	10	0.587	0.052	0.503	0.687
Toronto Raptors	10	0.497	0.122	0.331	0.685
Utah Jazz	10	0.55	0.085	0.447	0.732
Washington Wizards	10	0.54	0.077	0.411	0.652

Tables 3 and 4

Year	Observations	Mean	Std. Dev.	Min	Max
2009	30	0.497	0.053	0.378	0.605
2010	30	0.509	0.067	0.38	0.687
2011	30	0.503	0.092	0.215	0.636
2012	30	0.529	0.085	0.294	0.65
2013	30	0.547	0.085	0.332	0.7
2014	30	0.539	0.082	0.331	0.685
2015	30	0.544	0.066	0.421	0.656
2016	30	0.588	0.074	0.436	0.732
2017	30	0.607	0.06	0.458	0.729
2018	30	0.607	0.06	0.503	0.743

Conference	Observations	Mean	Std. Dev.	Min	Max
East	150	0.549	0.077	0.331	0.729
West	150	0.544	0.087	0.215	0.743

The Minnesota Timberwolves had the minimum Gini coefficient (.215), more than four standard deviations away from the initial mean; showing them to have a relatively more equal pay structure. In 2011, the Minnesota Timberwolves highest paid player was Michael Beasley making $6,262,347 and the player Malcolm Lee was paid the minimum wage at $500,000. Although this seems like a great difference, the rest of the team's pay structure was more cohesive. The second lowest paid player was Wayne Ellington at $1,154,040, which is six times less than Beasley. There are 9 out of 15 players making between $7 million and $4 million dollars, making the dispersion more cohesive.

When looking at Gini statistics by year, there seems to be a constant increase over time. This means that each year the pay dispersion has gotten progressively more unequal. Looking at the Gini coefficient through the years I was able to see if there were any years where there was more even pay dispersion or large income gaps that could potentially be explained.

There is also limited variance for the Gini statistics by team. The teams did not show much change of their average Gini coefficient for all ten seasons, as seen in Table 2.

Even though there are limited differences between each team's average Gini coefficients, there is slight variation between East and West conferences, as seen in Table 5. Both have high Gini coefficients showing large dispersion, the Eastern conference has .549 and Western conference has .544, but their minimum Gini's are very different: the difference being .116. For the East, the Gini minimum was .331 while the West had a minimum of .215 showing the pay dispersion for the West is more evenly distributed than that of the East. I look at conferences because even if one conference is less competitive there are still eight teams that are going to be in the playoffs from both conferences competing for a championship. Regardless of pay structure each conference will have eight teams competing for the overall accomplishment of productivity, a championship.

Table 5

Team	Observations	Mean	Std. Dev.	Min	Max
Atlanta Hawks	10	45.6	12.66	24	68
Boston Celtics	10	51.6	12.955	25	66
Brooklyn Nets	10	31.1	13.788	12	52
Chicago Bulls	10	45.4	13.986	22	71
Cleveland Cavaliers	10	44.9	23.369	19	73
Charlotte Hornets	10	35.4	13.006	7	51
Dallas Mavericks	10	44.3	14.291	24	73
Denver Nuggets	10	45.2	10.952	30	61
Detroit Pistons	10	33.3	6.482	25	44
Golden State Warriors	10	59.1	24.351	23	88
Houston Rockets	10	52.5	12.989	34	76
Indiana Pacers	10	47.1	10.333	32	66
LA Clippers	10	49	12.009	29	63
LA Lakers	10	38.8	17.844	17	73
Memphis Grizzlies	10	40.5	14.285	15	70
Miami Heat	10	45.7	12.685	22	64
Milwaukee Bucks	10	54.8	15.404	37	82
Minnesota Timberwolves	10	28.9	10.837	15	48
New Orlens Pelicans	10	36.2	9.852	21	53
New York Knicks	10	33.1	12.432	17	60
Oklahoma City Thunder	10	56.9	8.787	45	69
Orlando Magic	10	36.1	15.574	20	69
Philadelphia 76ers	10	33.5	16.236	10	58
Pheonix Sun	10	33.6	14.347	19	64
Portland Trail Blazers	10	47.4	10.532	28	61
Sacramento Kings	10	62.7	10.242	48	78
San Antonio Spurs	10	28.7	4.945	22	39
Toronto Raptors	10	47.7	17.777	22	74
Utah Jazz	10	42.2	11.717	25	57
Washington Wizards	10	37.3	13.132	20	56

Tables 6 and 7

Year	Observations	Mean	Std. Dev.	Min	Max
2009	10	43.7	16.507	12	73
2010	10	43.633	16.334	17	73
2011	10	35.993	13.668	7	62
2012	10	43.8	16.243	20	82
2013	10	43.967	16.273	20	82
2014	10	43.7	16.818	16	83
2015	10	43.367	17.862	10	88
2016	10	43.366	14.733	20	83
2017	10	44.067	15.456	21	76
2018	10	43.733	15.516	17	74

Conference	Observations	Mean	Std. Dev.	Min	Max
East	150	41.16	15.62	7	82
West	150	44.75	16.05	15	88

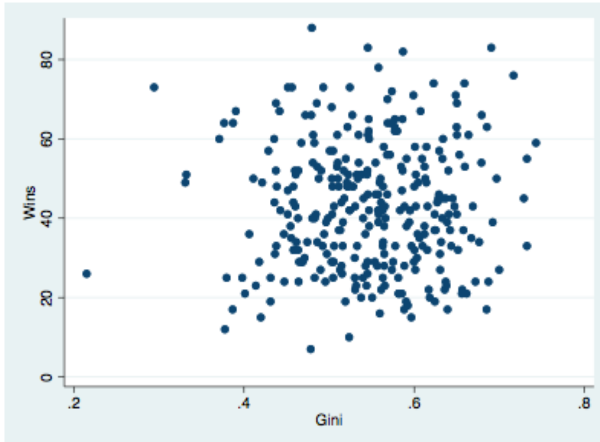
The average wins for the entire data set is 42.95 although the minimum is bringing this average down a bit (due to the low wins of the lockout year in 2011). The minimum wins for a team was 7 wins, this was the Cleveland Cavaliers, with a Gini of .478, showing great dispersion. Without the 2011 season, the minimum wins was 10, 3 more wins when including the 2011 season in the data set. The maximum wins was 88 in 2015, which was the performance of the Golden State Warriors with a Gini of .480, comparing not far off 2011 Charlotte Hornets Gini. This was also a year the Golden State Warriors made the playoffs and won the championship all with a Gini coefficient that depicts a large wage dispersion.

Methodology

In order to analyze the relationship between wage dispersion and productivity in the National Basketball Association, I start with a basic OLS model and then relax the following OLS concerns. The first concern, time dependence could be correlated with both dependent and independent variables, in particular, there may be autocorrelation between seasons. I address the autocorrelation by including a model that lags my dependent variable, *Wins*. In sports, an occurrence in one year will sometimes take time to show: the year before has something to do with the year after. The second concern is the presence of time invariant unobservables in team behavior. A potential confounding time invariant unobservable is team ownership or management philosophy, I control for these unobservables by implementing an FE model. The third concern, are time varying unobservables which I will be controlling for using an RE model. A potential time varying unobservable is locker room dynamics, throughout the year the locker room dynamic can change game to game. Although all of these models independently control for these concerns alone, in my analysis, incorporating both FE lagged and RE lagged models will each tackle these concerns simultaneously.

<u>6 OLS Assumptions</u>

OLS regression is based on six key assumptions that if violated mean our estimators are not the best linear unbiased estimators. The first assumption is that the relationship between independent and dependent variables is linear.



The second assumption states that no combination of independent variables has a linear relationship. In my data this means no variable is a scaling of another. For example, the East and West conferences are not their own independent variables, they are counted only once with *EastWest*. The third assumption is the expected value of errors should equal 0. If this does not hold true there is information in the errors that is not reflected in the model. Next, the fourth assumption is that all errors are independent of one another with constant variance. Knowing the value of one error will not tell you anything about the others. The fifth assumption claims that independent variables are non-stochastic, they are fixed or predetermined. This is important for any measurement error, in this case the potential violation due to the scale issue in salaries' not accounting for incentive money or bonuses could have potential to cause bias. Lastly assumption six, errors are normally distributed, this is important when making decisions about variables which are on the border of statistical significance.

Basic OLS Model

The first regression uses a basic OLS model; this will be the baseline. This model assumes that any unobservable factors affecting each team's performance are independent across each season and are uncorrelated with wage disparity. It is not clear that these are reasonable assumptions in the context of the NBA.

$$Wins = \alpha + \beta_1 Gini_{it} + \beta_2 EastWest_i + \beta_3 Year_{it} + \varepsilon_{it}$$

As earlier explained, scenarios in sports take time to play out. If a team wins the championship or makes playoffs they may attract different types of players to their organization who are willing to be paid less. They may also be more likely to win next season. For example, Kevin Durant moved to the Golden State Warriors in 2016 to play with "the trio" Steph Curry, Klay Thompson, and Draymond Green. According to Bleacher Report, a basketball supersite, there were five teams on the hunt for Durant, but only one team who had won the championship in 2015, the Golden State Warriors. Durant's addition to the Golden State roster will both drive down wage inequality because there will be more great players paid around the same amount, as well as creating changes in productivity with the addition of Durant to compliment "the trio". This, which both affects Gini coefficient and team productivity, is not independent over time. This evidence shows potential for assumptions to be violated in our baseline OLS model.

Another issue with this OLS model is that unobservables may be correlated with both our dependent and independent variables. The baseline OLS model assumes that the unobservables that affect team productivity are uncorrelated with Gini coefficient. An example of team fixed effects is management philosophies. Management has final say on many decisions which affect both salary and productivity (signing good team players even if it means going over soft salary cap and taking fines). They have the ability to create a culture with the decisions they make and the values they put into the organization. If they are more concerned about generating revenue than getting players who will help their team, even when they are no longer in that organization their philosophies live on. An example of team random effects is locker room dynamics. As a basketball player, from experience, it is always easy to be upbeat when things in your season are going in a positive direction, but this is not always the case when your team is facing tough challenges. These unobservable factors could potentially lead to bias in our OLS estimates. The following regressions control for these time dependent and unobservable factors.

Lagged Model

$$Wins_{it} = \alpha + \beta_1 Wins_{t-1} + \beta_2 Gini_{it} + \beta_3 EastWest_i + \varepsilon_{it}$$

The second regression includes lagged wins because wins last year are an important omitted variable that are correlated with both productivity and the Gini coefficient. For example, consider the previously stated (p. 25) example of Kevin Durant going to the Golden State Warriors to be with a winning team and to play in a winning environment. There are players whose contracts last more than one year and their salary will have something to do with what players can be brought in the year after. Their salaries will have some effect on how much room is left in the soft salary cap.

In this regression it is important to note that by trying to control for time dependence we fall into the failure of strict exogeneity: the error term is potentially correlated with our lagged wins variable. Therefore the results of this regression only hold true if the assumptions are met.

The Fixed Effects model

After addressing the lag regression it was important to address both time invariant and time varying unobservables.

$$Wins_{it} = \alpha + \beta_1 Gini_{it} + \beta_2 EastWest_i + \mu_i + \varepsilon_{it}$$

The third regression is the fixed effects (FE) model, to control for time invariant unobservable factors that could have a potential effect on team productivity. Fixed effects differentiate present year from the year before, removing all constants. As noted in the equation above, the fixed effects μ_i only has a subscript of *i,* this is because the *t* subscript is differentiated away, while the error term ε_{it}contains both *i* and *t* subscripts. This error term is the independently identically distributed error (IID).

Owners/managers and their philosophies within sports teams rarely change, making the decisions they make both impactful on productivity and salary dispersion. For example, the New York Knicks' owner James Dolon. Since Dolon became the owner of the Knicks in 1995 their productivity decreased greatly. His tenure is amongst the least successful owners in the modern history of the NBA. Good owners and managers have potential to make a team more productive by attracting better players or "superstars", by taking on luxury tax in order to obtain better players. There is also the superstar effect tradeoff explained in Mondello and Maxcy's (2009) NFL study where they found greater wage disparity is positively associated with franchise revenue. In this case, the player would be there as a public figure not as a team producing player. Decisions like this impact the environment of the organization, management philosophies live on, and if a front-office manager makes decisions based on revenue and not team productivity this could have an effect on team production. Anecdotally, the 2012 San Antonio Spurs foster an environment for winning. There long time coach Gregg Popovich, and front office executive R.C. Buford, emphasize team basketball. In that 2012 season, the Spurs led the league in team average assists with 25 per game, while also making it to the NBA Finals.

<u>The Random Effects model</u>

$$Wins_{it} = \alpha + \beta_1 Wins_{t-1} + \beta_2 Gini_{it} + \beta_3 EastWest_i + \mu_i + \varepsilon_{it}$$

Lastly, the fourth regression is the Random Effects (RE) model. This regression is important to control for time varying unobservable factors that could be correlated with team productivity and Gini coefficient. A time varying unobservable could be locker room dynamics. Anecdotally, the 2018 Houston Rockets starting guard Chris Paul was off to a great start of his season until he and the Rockets met the Lakers. During the game against the Lakers Paul got into a physical altercation with rival guard Rajon Rondo. This fight resulted in his ejection, as well as a two game suspension. This change in dynamics could have a negative effect on the Houston locker room. Locker room dynamics can change at any point in the season and are unquantifiable.

It is important to note that in this regression, as noted in the equation above, the random effects $\mu_{\square\square}$ has a subscript of *i* and *t,* the error term $\varepsilon_{\square\square}$ contains both *i* and *t* subscripts. As a consequence of containing the subscript *t,* the random effects models residuals are correlated across time (Angrist and Pischke p.166).

Fixed Effects with a Lag model

$$Wins_{it} = \alpha + \beta_1 Wins_{t-1} + \beta_2 Gini_{it} + \beta_3 EastWest_i + \mu_i + \varepsilon_{it}$$

To control for time dependence and time invariant unobservables it was imperative to incorporate both the lagged model and the fixed effects model simultaneously. This model controls for both, in our case, the previous years results having an effect on the present years results and management philosophies that live on in sports franchises. The problem with this FE Lag model regression is the potential for dynamic panel data bias, (Nickel 1981). The FE model differentiates itself over time, differentiating the year before from the present year, and the lagged wins variable crosses time constraints (looking at the previous year wins). The differenced residual is likely to be correlated with the lagged dependent variable (Angrist and Pischke p. 245). It is important to note that the results from this regression are only true if the OLS assumptions are met.

Random Effects with Lag

$$Wins_{it} = \alpha + \beta_1 Wins_{t-1} + \beta_2 Gini_{it} + \beta_3 EastWest_i + \mu_{it} + \varepsilon_{it}$$

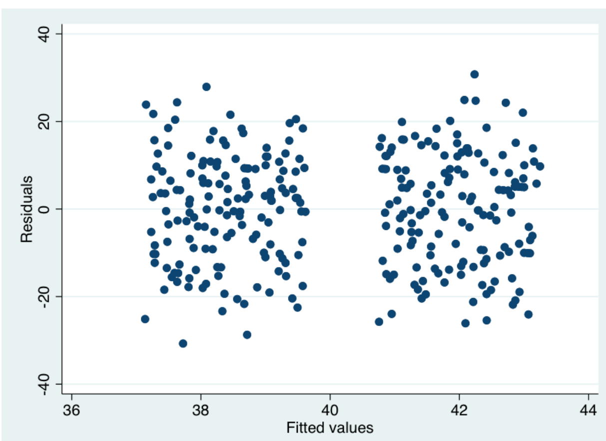
Similar to the FE Lag model, I also look at time dependence and time varying issues at the same time by running a random effects model with the lagged dependent variable wins. But just like the FE model, I run into potential dynamic panel data bias when lagging the dependent variable. This is because there could be correlation between μ_{it} and $Wins_{it}$ and $Gini_{it}$.

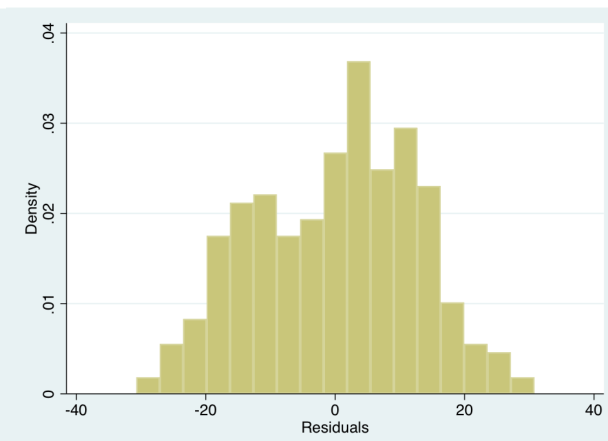
Results

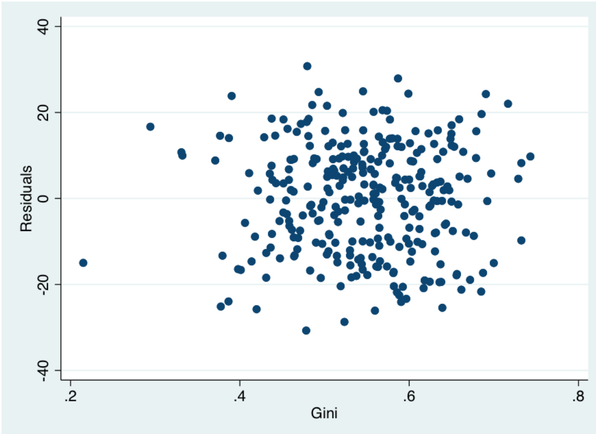
Regression Model	Coefficient	T-Value	P-Value	Significance?
Basic OLS	1.174	0.12	0.606	Insignificant
Lagged	-3.169	-0.43	0.669	Insignificant
FE	9.589	1.05	0.293	Insignificant
FE with Lag	1.741	0.2	0.842	Insignificant
RE	8.194	0.94	0.347	Insignificant
RE with Lag	-7.819	-0.96	0.337	Insignificant

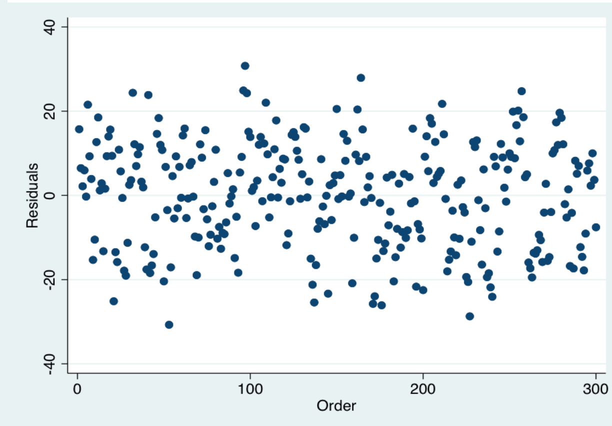
Our baseline OLS model finds no statistically significant relationship between a *Gini* and *Wins*. The Gini had a p-value of .606, and a positive coefficient of 1.174: a positive correlation but is not statistically significant.

Although Gini did not have any statistical significance, the variable *Eastwest*, the conference dummy variable, did have statistical significance at the .05 level. I estimate that, all else held equal, being in the Western conference on average predicts 3 more wins per season. This could have some significance to wins because since 2009 the West have had seven championships while the East have had only four (NBA History Recaps, 2019). These basic OLS results are only suggestive because the assumptions do not hold true.









When looking at the residual analyses the residual versus fitted plot seems to have a pattern of two distinct clusters. This is explained by the fundamental principle that if one team wins another must lose, driving the two clusters apart. The histogram of the residuals shows a reasonably normal distribution, showing a univariate graph. The residuals versus Gini coefficient scatter plot showed no pattern which means there was no relationship between the error and the Gini coefficient. But there is a potential outlier, this outlier is the Minnesota Timberwolves with a small Gini of .215 with only 7 wins on the season. There is also no pattern in the residuals versus order scatter plot. The one plot that does draw attention is the residual versus fitted plot, showing there may be a correlated omitted variable.

To control for time independence, my second regression uses a lagged model and finds no evidence of a relationship between *Gini* and *Wins*. We include wins from last year as a regressor, meaning there are only 270 observations as opposed to 300. The Gini coefficient's p-value was .669 and its coefficient was -3.169, showing a statistically insignificant negative relationship between pay dispersion and wins.

There was statistical significance in this regression for current wins and lagged wins: every additional win last year predicts .591 additional present years wins. This model suggests a lack of strict exogeneity - lack of independence between independent variables and dependent variables, this relationship was expected. Despite this bias, the model still faces the flaw of not controlling for time varying and invariant unobservables.

When looking at the residual analyses, the residuals versus fitted plot shows there is no relationship between the error and predicted wins. There are no longer two clustered patterns that are shown in the Basic OLS residual versus fitted plot. Suggesting, in the Basic OLS regression there is a correlated omitted variable (lagged wins) and we are attributing too much of the results to the Gini coefficient. The histogram of the residuals follows a normal distribution but had a slight left tail. The residuals versus Gini scatter plot still has no pattern and still had the low Gini outlier (2011 Minnesota Timberwolves). The residuals versus order is very scattered with absolutely no pattern, showing no relationship between the errors and the *Gini*. Although this regression controls for time dependence, it does not address the potential of unobservables and the assumptions are not likely to be true.

The fixed effects model (FE) controls for time invariant unobservable factors. In this regression there is still no significant relationship between the Gini coefficient and *Wins*. *Gini* had a p-value of .293 and a coefficient of 9.589, and these results do not control for time dependence.

To incorporate both lagged and FE models, I run the FE model with the lag as a covariate. The results showed lagged wins were again statistically significant, each additional win last year can account for .408 of a win in the present year. The Gini coefficient did not have any statistical significance with p-value of .842 and a coefficient of 1.741, but I acknowledge the likeliness for dynamic panel data bias (Nickel 1981).

Lastly, the fourth regression is the Random Effects (RE) model controls for time varying unobservable factors. The RE model found no significant relationship between Gini coefficient *Wins* with a p-value of .347 and a coefficient of 8.194.

Just like the FE model, the RE model alone does not control for time dependence. The RE model must be paired with the lagged model to control for both concerns. Just like the FE model, I lagged wins to see if there was any relation to team productivity. Similarly, to the FE lagged model in the RE lagged model there are statistically significant factors that could affect team productivity, each additional win last year can account for .408 of a win this year. The Gini coefficient does not have any statistically significant relationship on team productivity. It is important to note, like the FE Lag model, these results are suggestive and there is potential for dynamic panel data bias (Nickel 1981).

Interpretation

Through all these regressions the Gini coefficient is not statistically significant. The Gini coefficient's p-values were all above 0.05 as seen in above (p. 30). When adding lagged wins to the regression and controlling for time dependence, there were results, but the variable in question, *Gini,* was not significant. Even when controlling for time varying and time invariant unobservables there was no relationship between the *Gini* and the dependent variable. When combining the regressions, to control for both issues, creating an FE Lag and the RE Lag there was still no effect on team productivity. These estimators in each regression are only true if the assumptions hold. As a basketball player, I know it is impossible to not control for these occurrences, previous year records, coaching/organization strategies, and locker room dynamics all play a huge role in the productivity of a team. Given persistent insignificance on the variable of interest I do not conduct the Blundell and Bond dynamic panel data estimator to address potential dynamic panel data bias.

Conclusion

I began this study convinced that cohesion theory would in fact describe and predict the effect of inequality on team productivity. Results show that there is no statistically significant relationship between the wage dispersion and team productivity at the .05 level. In each model, basic OLS, lagged, FE and RE the Gini coefficient had insignificant effects.

These results showed that between 2009-2018 there was no significant effect of wage dispersion on team wins within the NBA when including year and conference in the regressions. Using the conferences variable *EastWest* was important due to the way the playoffs in the NBA are structured and how one team can obtain more wins by going further in playoffs against their own conference opponents. It was important to use all four models to find the best linear unbiased model to depict the relationship. There was no relation between Gini and *Wins* in the baseline OLS model. After this baseline doing a lagged model, lagging wins because the wins of the previous year may have some effect on this present year's results. It is shown that Gini had no effect on wins but lagged wins, suggesting the correlated omitted variable, does have an effect. But still with the lagged regression there is a no control for time variant and invariant unobservables. These are controlled for by the FE and RE models respectively. Both of which also showed no statistical significance between Gini coefficient and wins. Using a combined FE model and lagged wins we found no statistically significant results to conclude that the Gini coefficient has an effect on a teams productivity. Using a combined RE model and lagged wins there was also no significance of the variable of interest. Although as noted there is potential for dynamic panel data bias in both of these regressions. These results are only true if the assumptions are correct.

Through these results, either both cohesion theory and tournament theory are working simultaneously. Initially as a basketball player, I hypothesized that cohesion theory would hold true, but I was also curious as to how money could have an effect on professional athletes in the NBA – making them follow cohesion theory. Since this study was at the macro level of the entire NBA it is possible that some teams benefitted while others did not benefit from pay inequality, which summed out to insignificant effects overall. Future avenues of research could include seeing if these findings hold true for the top and bottom tier teams in the NBA going back further historically. If anyone holds the information for bonuses and incentive money I would hope they could recreate this study including that data. Lastly, another study could be looking at the NBA by conference and see if either of these conferences characterize either theory.

And finally, while my results show that pay inequality does not negatively affect productivity as I initially hypothesized, it is no less significant to discover that inequality does not positively affect productivity. This has important consequences for teams that feel that have no choice but to go for the 'superstar effect' if they are to win championships.

Appendices

Basic OLS Regression

```
reg wins gini eastwest year

      Source |       SS           df       MS      Number of obs   =       300
-------------+----------------------------------   F(3, 296)       =      2.35
       Model |  1110.86229         3  370.287431   Prob > F        =    0.0728
    Residual |  46694.7744       296  157.752616   R-squared       =    0.0232
-------------+----------------------------------   Adj R-squared   =    0.0133
       Total |  47805.6367       299  159.885072   Root MSE        =     12.56

------------------------------------------------------------------------------
        wins |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        gini |   1.174083   9.905113     0.12   0.906    -18.31929    20.66745
    eastwest |   3.572265   1.451068     2.46   0.014      .716547    6.427982
        year |   .2347037   .2831473     0.83   0.408    -.3225332    .7919405
       _cons |  -434.8274   567.6752    -0.77   0.444    -1552.018    682.3635
------------------------------------------------------------------------------
```

OLS with a Lag Regression

```
. reg wins l.wins gini eastwest

      Source |       SS           df       MS      Number of obs   =       270
-------------+----------------------------------   F(3, 266)       =     51.93
       Model |  15726.5578         3  5242.18592   Prob > F        =    0.0000
    Residual |  26853.3089       266  100.952289   R-squared       =    0.3693
-------------+----------------------------------   Adj R-squared   =    0.3622
       Total |  42579.8667       269  158.289467   Root MSE        =    10.048

------------------------------------------------------------------------------
        wins |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wins |
         L1. |   .5905731   .0487918    12.10   0.000     .4945059    .6866403
             |
        gini |  -3.169165   7.400801    -0.43   0.669    -17.74077    11.40244
    eastwest |    1.52227   1.236466     1.23   0.219     -.912235    3.956775
       _cons |   17.40553   4.449059     3.91   0.000     8.645675    26.16538
------------------------------------------------------------------------------
```

FE Regression

```
. xtreg wins gini eastwest , fe
note: eastwest omitted because of collinearity

Fixed-effects (within) regression               Number of obs     =        300
Group variable: team_id                         Number of groups  =         30

R-sq:                                           Obs per group:
     within  = 0.0041                                         min =         10
     between = 0.0035                                         avg =       10.0
     overall = 0.0008                                         max =         10

                                                F(1,269)          =       1.11
corr(u_i, Xb)  = -0.0617                        Prob > F          =     0.2933

------------------------------------------------------------------------------
        wins |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        gini |   9.589414   9.106446     1.05   0.293    -8.339557    27.51839
    eastwest |          0  (omitted)
       _cons |   34.93186   5.021145     6.96   0.000     25.04612     44.8176
-------------+----------------------------------------------------------------
     sigma_u |  7.2260393
     sigma_e |  11.022783
         rho |  .30057811   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(29, 269) = 4.28                     Prob > F = 0.0000
```

FE with Lag Regression

```
Fixed-effects (within) regression               Number of obs     =        270
Group variable: team_id                         Number of groups  =         30

R-sq:                                           Obs per group:
     within  = 0.1745                                         min =          9
     between = 0.9250                                         avg =        9.0
     overall = 0.3641                                         max =          9

                                                F(2,238)          =      25.15
corr(u_i, Xb)  = 0.5194                         Prob > F          =     0.0000

------------------------------------------------------------------------------
        wins |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wins |
         L1. |   .4081175   .0585065     6.98   0.000     .2928608    .5233741
             |
        gini |   1.741043   8.748366     0.20   0.842    -15.49308    18.97516
    eastwest |          0  (omitted)
       _cons |   22.76747   5.073702     4.49   0.000     12.77238    32.76257
-------------+----------------------------------------------------------------
     sigma_u |  4.6574164
     sigma_e |  9.8180635
         rho |  .18369278   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(29, 238) = 1.45                     Prob > F = 0.0693
```

RE Regression

```
. xtreg wins gini eastwest , re

Random-effects GLS regression                   Number of obs      =       300
Group variable: team_id                         Number of groups   =        30

R-sq:                                           Obs per group:
     within  = 0.0041                                         min =        10
     between = 0.0570                                         avg =      10.0
     overall = 0.0206                                         max =        10

                                                Wald chi2(2)       =      2.73
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.2556

------------------------------------------------------------------------------
        wins |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        gini |   8.194332   8.715437     0.94   0.347    -8.887611    25.27627
    eastwest |   3.605737   2.626711     1.37   0.170    -1.542523    8.753996
       _cons |   33.89202   5.135152     6.60   0.000      23.8273    43.95673
-------------+----------------------------------------------------------------
     sigma_u |  6.3014675
     sigma_e |  11.022783
         rho |  .24631496   (fraction of variance due to u_i)
```

RE with Lag Regression

```
Random-effects GLS regression                   Number of obs      =       270
Group variable: team_id                         Number of groups   =        30

R-sq:                                           Obs per group:
     within  = 0.1738                                         min =         9
     between = 0.9062                                         avg =       9.0
     overall = 0.3693                                         max =         9

                                                Wald chi2(3)       =    155.78
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
        wins |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        wins |
         L1. |   .5905731   .0487918    12.10   0.000      .494943    .6862032
             |
        gini |  -3.169165   7.400801    -0.43   0.668    -17.67447    11.33614
    eastwest |    1.52227   1.236466     1.23   0.218    -.9011583    3.945698
       _cons |   17.40553   4.449059     3.91   0.000     8.685532    26.12552
-------------+----------------------------------------------------------------
     sigma_u |          0
     sigma_e |  9.8180635
         rho |          0   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Bibliography

Adcroft, Andy, et al. "The impact of salary dispersion and performance bonuses in NFL organizations." *Management Decision* (2009).

Angrist, Joshua D., and Jörn-Steffen Pischke. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press, 2008.

Annala, Christopher N., and Jason Winfree. "Salary distribution and team performance in Major League Baseball." *Sport Management Review* 14.2 (2011): 167-175.

Berri, David J., and R. Todd Jewell. "Wage inequality and firm performance: Professional basketball's natural experiment." *Atlantic Economic Journal* 32.2 (2004): 130-139.

Coates, Dennis, Bernd Frick, and Todd Jewell. "Superstar salaries and soccer success: The impact of designated players in Major League Soccer." *Journal of Sports Economics* 17.7 (2016): 716-735.

Ding, Daniel Z., Syed Akhtar, and Gloria L. Ge. "Effects of inter-and intra-hierarchy wage dispersions on firm performance in Chinese enterprises." *The International Journal of Human Resource Management* 20.11 (2009): 2370-2381.

Katayama, Hajime, and Hudan Nuch. "A game-level analysis of salary dispersion and team performance in the National Basketball Association." *Applied Economics* 43.10 (2011): 1193-1207.

Lawler, E. E. "Ⅲ.(1971). Pay and organizational effectiveness."

Levine, David I. "Cohesiveness, productivity, and wage dispersion." *Journal of Economic Behavior & Organization* 15.2 (1991): 237-255.

“NBA G-League 101: Two Way Contracts”, NBA G League,NBA Media Ventures, https://gleague.nba.com/nba-g-league-101-two-way-contracts/

"NBA Playoffs Bracket 2019: Full Schedule, Dates, Times, TV Channels For Conference Finals". *Sportingnews.Com*, 2020, https://www.sportingnews.com/us/nba/news/nba-playoffs-bracket-2019-schedule-dates-times-tv-channels-conference-finals/1sjww7k5mndy2158vb85n2eot3.

NBA Season Recaps: 1946-2019, the NBA.com staff, 2019. https://www.nba.com/history/season-recap-index

Shaw, Jason D., Nina Gupta, and John E. Delery. "Pay dispersion and workforce performance: Moderating effects of incentives and interdependence." *Strategic Management Journal* 23.6 (2002): 491-512.

Shaw, Jason D. "Pay dispersion." *Annu. Rev. Organ. Psychol. Organ. Behav.* 1.1 (2014): 521-544.

Slottje, Daniel J. "Using grouped data for constructing inequality indices: Parametric vs. non-parametric methods." *Economics Letters* 32.2 (1990): 193-197.

"Teams", NBA and WarnerMedia News & Sports, https://www.nba.com/teams

USA Today Sports HoopsHype

LandOfBasketball

Basketball-Reference