

PRODUCTIVITY AND INTEREST ARBITRATION IN
MAJOR LEAGUE BASEBALL:
3 ESSAYS

A Dissertation
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

by

Jonathan R. Brown
May 2017

Examining Committee Members:

Dr. Michael Leeds, Advisor, Department of Economics
Dr. Joel Maxcy, Drexel University, Department of Sports Mangement
Dr. Catherine Maclean, Department of Economics
Dr. Brantly Callaway, Department of Economics
Dr. Joris Drayer, Department of Sport and Recreation Management

©
Copyright
2017

by

Jonathan R. Brown

All Rights Reserved

ABSTRACT

This dissertation consists of three essays exploring the labor market in Major League Baseball (MLB), as well as the negotiation environment for arbitration eligible players. Chapter 1 will show that the distribution of individual labor productivity has a significant effect on overall firm output. Results indicate that a firm with heterogeneous workers should consider not only the sum of individual contributions, but also how individual contributions are allocated, as increased concentration reduces overall output. Traditionally, salary dispersion acts as a proxy for ability dispersion. Past literature indicates that workers respond to disparity, though the literature is conflicting as to the direction of this response. In all cases, however, using salary as a proxy for worker production over-simplifies the firm's decision-making process. This chapter uses data from MLB to measure productivity concentration directly, independent of wage concentration. If workers act as complements to one another then the concentration of productivity will influence overall output in a way that is unrelated to the distribution of salaries. This analysis allows for workers' behavioral response to wage, and the productivity effects of heterogeneous individual production levels, to be evaluated separately. The analysis could extend to any industry in which workers act as a team contributing a portion of a final product. These findings are particularly useful for industries, like MLB, in which varying degrees of monopsonistic power make wage a poor proxy for productivity.

Chapter 2 discusses Final Offer Interest Arbitration (FOA), a bargaining mechanism designed to promote private negotiation and, when necessary, resolve a negotiation impasse without a work stoppage. If parties cannot settle on mutually acceptable terms, they bring their proposed terms to an arbitrator. The arbitrator then rules in favor of one party. In a Final Offer system, the winning party's terms become the binding terms of the agreement. In order for FOA to be an effective mechanism, it should promote bargaining, meaning it is used relatively infrequently and, when used, its outcomes should resemble privately negotiated terms. Traditionally, both parties are given equal power to select the arbitrator who will hear the case. This veto power during the selection process should weed out any calculable or predictable favoring of one party over another, so an FOA system should not yield significantly different settlements from those that do not go to arbitration and are instead privately negotiated. This chapter explores the use of FOA in Major League Baseball. Different players face arbitration eligibility at points during their career, allowing for a side-by-side view of settlements with and without an FOA mechanism. Results indicate that FOA has succeeded in promoting bargaining, but that a bias against players lingers even as uncertainty dwindles.

Chapter 3 is meant to complement chapter 2 by further exploring FOA. Arbitrators must maintain a degree of unpredictability in order to promote private negotiation. However, they also must be predictable enough that parties expect a "high-quality" ruling, meaning the outcome falls within a range that parties believe reflects privately negotiated decisions. I use data from Major League Baseball (MLB) and the Wisconsin Employment Relations Commission (WERC) to explore the way teams learn from past decisions to reduce uncertainty surrounding arbitrator decisions.

TABLE OF CONTENTS

ABSTRACT	iii
LIST OF TABLES	viii
LIST OF FIGURES	x

CHAPTER

1 INDIVIDUAL LABOR PRODUCTIVITY CONCENTRATION IN MAJOR LEAGUE BASEBALL	
1.1 Introduction	1
1.2 Related Literature	4
1.3 Conceptual Framework	8
1.4 Empirical Model, Variables, & Data	11
1.4.1 <i>Empirical Model</i>	11
1.4.2 <i>Variables</i>	15
1.4.3 <i>Data</i>	22
1.5 Results	22
1.6 Discussion & Conclusion	27
1.7 Future Work	29
2 FIRM BEHAVIOR UNDER CONDITIONS OF FINAL OF- FER ARBITRATION: AN EMPIRICAL ANALYSIS OF MA- JOR LEAGUE BASEBALL'S SALARY NEGOTIATION STRUC- TURE	

2.1	Introduction	30
2.2	Institutional Background	34
2.3	Related Literature	37
2.4	Conceptual Framework	41
2.5	Data and Variables	49
2.5.1	<i>Data</i>	49
2.5.2	<i>Variables</i>	49
2.6	Empirical Analysis	52
2.6.1	<i>OLS and Fixed Effects Model</i>	52
2.6.2	<i>OLS and Fixed Effects Results</i>	54
2.6.3	<i>Quantile Regression Model</i>	59
2.6.4	<i>Quantile Regression Results</i>	62
2.7	Discussion and Conclusion	68
3	ARBITRATION AND UNCERTAINTY : AN EMPIRICAL ANALYSIS OF MAJOR LEAGUE BASEBALL'S NEGOTIATION LEARNING PROCESS	
3.1	Introduction	74
3.2	Related Literature	76
3.3	Data	80
3.4	Conceptual Framework	84
3.5	Empirical Model	87
3.6	Results	88
3.7	The Player's Perspective	91
3.8	Comparing Major League Baseball and the Public Sector	94

3.9 Discussion and Conclusion	98
BIBLIOGRAPHY	99
APPENDICES	
A CHAPTER 1-LORENZ CURVES	105
B CHAPTER 1-ROBUSTNESS CHECK	106

LIST OF TABLES

TABLE	Page
1.1 OLS Regression of $W_{i,t}$ on WARP	18
1.2 Cross-Correlation Table - Team Data 1974-2014	19
1.3 Summary Statistics	21
1.4 Estimation with Direct Productivity Measures	23
1.5 Variance Inflation Factor	24
1.6 Estimation with Direct Productivity and Salary Measures	25
1.7 Estimation without Direct Productivity Measures	26
2.1 Summary Statistics	52
2.2 Pooled OLS and Fixed Effects estimations, Factors Affecting Average Yearly Salary of Arbitration Eligible Players	54
2.3 Contracts Removed from Sample as Robustness Check	57
2.4 Pooled OLS and Fixed Effects estimations with 4+ year Contracts Removed, Factors Affecting Average Yearly Salary of Arbitration Eligible Players	58
2.5 Quartile Regression Results, Factors Affecting Average Yearly Salary of Arbitration Eligible Players.	62
2.6 Decile Regression Results (A), Factors Affecting Average Yearly Salary of Arbitration Eligible Players	63

2.7	Decile Regression Results (B), Factors Affecting Average Yearly Salary of Arbitration Eligible Players	64
3.1	Summary Statistics	82
3.2	Fixed Effects and Arellano Bond Estimation of Total Disagreement Factor	88
3.3	Fixed Effects and Arellano Bond Estimation of Average Disagreement Factor	91
3.4	Probit Estimation of Arbitration Hearing Outcomes	93
3.5	Fixed Effects Estimator Controlling for Macroeconomic and Political Factors in Public Sector Arbitration Hearing Outcomes	96
B.1	Pooled, Fixed Effects, and Arellano-Bond Estimation with and without Salary Measures - Robustness Check	107

LIST OF FIGURES

FIGURE	Page
3.1 Total MLB Hearings 1974-2014	84
3.2 Total WERC Hearings 1972-2010	94
A.1 Lorenz Curve Comparison	105

CHAPTER 1

INDIVIDUAL LABOR PRODUCTIVITY CONCENTRATION IN MAJOR LEAGUE BASEBALL

1.1 Introduction

In a variety of situations, firm output cannot be characterized as the sum total of individual worker output. Instead, there is a conversion process wherein specialized individual contributions are combined and converted to firm-level output. In these cases, worker output is measured on a different scale than firm output. Workers are judged by their ability to efficiently and effectively fulfill their specific role. It is up to the firm to orchestrate an environment in which these individual contributions are optimally combined to generate firm-level output. If the firm cannot convert from individual to firm-level output, then a portion of individual efforts may never yield firm-level output, and are thus “lost in translation”. Since workers are paid for fulfilling their role regardless of the firm’s ability to make use of their contribution, the firm should target a distribution of labor talent that minimizes wasted inputs. This paper uses both salary and production data from Major League Baseball (MLB) to test the hypothesis that a team’s success is driven by both the overall level of individual worker inputs, and the dis-

tribution of individual worker inputs. Additionally, I contribute to the literature by isolating and separating the production effects of disparate talent levels working together from the behavioral effects of workers' response to their environment beyond the direct production process. By using the metric Wins Above Replacement Player, a proprietary statistic developed by *Baseball Prospectus*, I am able to calculate a Herfindahl-Hirschman index (HHI) to measure the disparity within each team's total individual contributions. I do the same to measure the distribution of salary. Results show that a team's ability to convert individual effort into firm-level output is directly related to the distribution of individual productivity. Specifically, as individual contributions become concentrated, teams become less efficient in the conversion process. MLB data is attractive for this analysis, as the measurements of productivity are easily categorized into individual contributions and team level output, and those measurements are well-defined and publicly available. I use both a team fixed effects and Arellano-Bond estimator and find that difference in production concentration and salary concentration can each explain a variation of over 6 wins. Past literature traditionally uses salary as a proxy for production. By analyzing production concentration and salary concentration separately, I am able to show that this literature is most likely suffering from omitted variable bias. MLB is categorized by a very specific production process, but the results should be of interest to any firm wherein individual workers fill specialized roles to contribute portions of a larger firm-level output. This includes any industry in which output is produced by a team of individuals with potentially heterogeneous productivity levels. The theoretical model suggests that the extent to which MLB results apply to other industries will depend on the degree to which inputs act as complements.

Further, this paper views productivity effects and behavioral effects as separate considerations rather than using salary as a proxy for ability, as is

standard in the literature. This distinction is important, as it has been well-documented that MLB's monopsony power can be exercised to varying extents depending on a player's tenure in the league, which makes salary a poor proxy for productivity. A monopsony will typically suppress wages, but in the case of MLB, this suppression is not non-uniform among players. Directly considering productivity allows teams to determine the distribution of talent that most efficiently converts individual contributions to team level output. For example, a team must consider whether it is better served by one highly productive worker, or two mid-level workers. A player's market value is determined by individual contributions regardless of how successful the team is. If teams are paying for individual output that is not being converted to team level output, then a portion of the wage bill is wasted. A firm can avoid paying for wasted effort by revising its talent distribution. The cost of the highly productive player vs. the cost of two mid-level players can and should be considered separately. By doing so, one is able to gain a view of productivity effects independent of salary, as well as a more accurate and isolated view of behavioral responses to salary disparity.

The current literature, discussed in section 1.2, thoroughly judges the impact of salary disparity on productivity; but oversteps its bounds when it draws conclusions about the impact of talent concentration based on analysis of salary concentration. Section 1.3 outlines a theoretical model for team production. Section 1.4 introduces the empirical model, variables and data used in the analysis. Section 1.5 shows results and outlines robustness checks. Section 1.6 discusses the analysis and offers concluding remarks. Section 1.7 outlines possible extensions and future work.

1.2 Related Literature

The most prominent theory in the wage concentration literature is that of cohesion, developed by Akerlof and Yellen (1988) and expanded by Levine (1991). Akerlof and Yellen build their model using the framework for efficiency wages laid out by Solow (1979), in which output is an increasing function of worker hours and effort expended. They further define effort as a function of the variance of wages paid by respective firms. Effort is decreasing in variance. As wage disparity grows, workers are more likely to feel that their wage is unfair relative to coworkers and thus exert less effort. This is an ad hoc assumption concerning the relationship between the distribution of wages and effort in the Solow Model. Akerlof and Yellen borrow from exchange theory in sociology and equity theory in psychology to build a microfoundation for their model. They define effort as a function of the ratio of the actual wage to the perceived fair wage, which they call the fair wage/effort hypothesis. This theory suggests that, as wage disparity grows, individual workers feel slighted and exert less effort, even when productivity levels are disparate. Production suffers as a result of these negative behavioral responses to wage. The less wage disparity that exists within a firm, the more effort workers exert. This creates an environment in which the worker cohesion necessary to generate firm-level output is optimized.

Alternative theories to the fair wage/effort hypothesis were introduced by Lazear and Rosen (1981) and Ramaswamy and Rowthorn (1991). While the authors offer slightly differing intuition to support their respective results, both papers suggest that, under certain conditions, designing compensation packages with disparities in wages is a useful tool when motivating workers to produce the efficient level of output. Lazear and Rosen approach their model from a tournament standpoint. The high wages of successful peers provides an incentive for workers to exert extra effort in an attempt to break into the higher wage bracket.

Ramaswamy and Rowthorn interpret the positive relationship they find between firm-level productivity and wage disparity via a “damage potential” framework. Firms allocate wage according to which employees could do the most damage to firm production by shirking. By paying the most to employees with the highest damage potential, firms can minimize the effects of shirking and therefore maximize productivity. These conflicting theories are of particular interest because, in both cases, the concentration of wages have a meaningful impact on firm output. It is the direction of the impact that is debated. This conflict provides a foundation for the my research question, by indicating that firms must consider the distribution of individual inputs. The competing hypotheses have not gone unexplored in the literature, though past studies have focused on salary as a proxy for productivity.

Professional sports are a popular setting for testing these theories since data are readily available on individual and team productivity as well as salary. Additionally, the self-contained world of sports lends itself to statistical inference. The controlled environment in which sports exists provides an excellent source of information with relatively limited endogeneity issues. Baseball has the most detailed data and is thus more frequently used, though similar analysis has been applied to the NFL, NBA, and NHL.

MLB data has been used extensively to support the notion that increased wage dispersion decreases productivity (Depken 2000, DeBrock, Hendricks, Koenker 2004, Annala and Winfree 2010, Breunig, Garrett-Rumba, Jardin, Rocaboy 2012). Mondello and Maxcy (2009) provide additional support for this theory using NFL data. I emulate Depken (2000) by using an HHI as a measure of wage concentration. Annala and Winfree (2010) expand Depken’s dataset to 2004 and use a Gini coefficient rather than the HHI. Their findings add empirical support for cohesion theory. My paper will primary follow in the path of Dep-

ken (2000) in the use of HHI as a measure of disparity, though I calculate Gini coefficients as a robustness check, similar to Annala and Winfree (2010).

Of particular importance to my analysis is Kamada and Katayama (2014) in which the authors use an Arellano-Bond estimator to correct for dynamic panel bias, a methodology from which I will borrow later on. Kamada and Katayama use data from Japan's Nippon Professional Baseball league (NPB), to test whether earlier results could be recreated outside of the United States (US). In contrast to earlier work, they find that wage concentration can improve team performance, though they do not suggest that their findings dispute past literature. Instead, they attribute their contrasting results to institutional differences between the United States and Japan. For example, Japan has lower wage variance in general, so average disparity is lower. Additionally, NPB players do not attain free agency as quickly as MLB players, so it is easier for NPB teams to suppress salaries. Institutional settings can dictate the direction of the causal relationship between wage disparity and production. Teams or firms may experience reactions to wage distribution in line with both sides of the conflicting theory, but their setting determines which effect dominates. This suggestion of simultaneous and conflicting effects would be analogous to theory's suggesting that a rise in real wage both increases and decreases an individual's labor hours supplied according to the substitution and income effect. Just as the baseline wage determined which of the substitution or income effect dominates in the decision making process, the institutional setting determines the net effect of wage disparity.

The literature commonly uses salary as a proxy for productivity. In a perfectly competitive labor market, wages equal each worker's marginal revenue product. This allows researchers to draw conclusions about the distribution of talent and productivity from the distribution of wages. However, various market failures make the wage an impractical proxy for individual productivity. In prac-

tice, earning a lion's share of the wage bill does not necessarily equate to providing a lion's share of the individual output. In the case of MLB, a team's monopsony power can be exerted to varying extents depending on a player's tenure in the league. This allows teams to suppress wage for certain groups of players. Data shows that this is the case in MLB, the NFL and the NBA as well (Krautmann, Berri, von Allmen 2009).

Workers choose a utility-maximizing level of effort in response to their compensation package. However, a worker may exert maximum effort, but still not be able to convert the effort into output. For this reason, productivity must be considered independent of wage.

Any attempt to capture the effect of salary on output without controlling for talent will most likely suffer from omitted variable bias. While salary may be a poor proxy for talent, the two variables are most likely correlated. Regressing firm-level productivity on salary alone inflates the partial effect. Breunig et al. (2012) anecdotally justify salary as a productivity proxy by pointing to Prince Fielder's 2011 contract with the Milwaukee Brewers. His \$15.5M contract, negotiated to avoid arbitration, made him one of the top-paid players on the team, despite his not yet reaching free agency. In this case, Fielder's wage should have been suppressed by the Brewers' monopsony power, as he was unable to entertain competing offers from any other team. Despite this market failure, the strong correlation between pay and performance was still realized. However, conflicting anecdotal evidence is just as easy to come by. Tim Lincecum made a combined total of \$1.55M for the 2008 and 2009 seasons, in which he won back to back Cy Young awards with the San Francisco Giants. Mike Trout of the Los Angeles Angels of Anaheim, in his first three full years in the major leagues, was named American League (AL) Rookie of the Year, AL MVP runner-up, and AL MVP, respectively. During those three years, he earned a combined total salary

of less than \$2M. These contracts were signed before Lincecum's and Trout's performance levels were revealed. However, the teams' monopsony power left both players unable to renegotiate terms even after revelation of their respective levels of productivity. In this case, the direct correlation between pay and productivity breaks down. To understand the partial effects of both talent and salary distribution on team success, both must be considered separately. To ignore one, or try to glean information on one from the other, introduces bias.

1.3 Conceptual Framework

Workers rarely produce output on their own. Rather, worker productivity is measured by how efficiently the worker performs his particular portion of the production process. If workers are perfect complements (as in a fixed proportion production function), then the level of output is determined by the least productive worker. Whenever labor inputs are complements in production to any degree, overall firm output is not necessarily the sum of each worker's individual output. Individual output that is successfully complemented will be converted to firm level output. Individual effort that is left over or unable to find the necessary complements is wasted. The distribution of performance plays an important role in the success of the firm, as it contributes to the relationship between individual inputs and firm output. As individual inputs become concentrated among a small number of workers, it becomes less likely that those productive workers will find sufficient complements from less-productive workers to collectively translate their inputs into output. In this case, a portion of the contributions from highly productive workers is wasted.

Scully (1974) uses this distinction between individual and team production to create the first empirical analysis of marginal revenue product. Scully acknowledges the problem posed by the lack of correlation between individual and

team output. In order to place a value on specific players, Scully has to separate individual contributions from team success, given that teams are often unable to convert individual contribution to team production. He looks at the correlation between slugging average (SLG), a measure of total bases taken per at-bat, and runs scored at the team level. He then generated an estimate of what a team would look like if a given player's SLG applied to the entire team. The rate of success for this hypothetical team could be applied to the percentage of plate appearances that the player actually took. This model formed valuations for players, independent of their teammates' production. In order to generate player MRP, Scully needed to work around the problem of individual contribution being wasted if inadequately complemented. This framework is extremely helpful in identifying the true value of individual players. However, from the firm side, this problem cannot be isolated and avoided. It must instead be considered directly. It is important for team to understand what contributes to wasted effort and how it affects overall output.

Specifically, teams should be concerned with the waste that stems from an unequal distribution of individual inputs. In order to test the hypothesis that increased concentration of individual inputs reduces team-level output, I first develop the following simple theoretical model. This framework links individual inputs to team output while looking at how the distribution of player talent affects production.

The model first outlines team output, measured in single-season wins for team j (W_j), as a function of runs produced ($r_{P,j}$) and runs allowed ($r_{A,j}$) such that

$$W_j = f(r_{P,j}, r_{A,j}) \tag{1.1}$$

W is increasing in r_P and decreasing in r_A . Since the only necessary step for winning a baseball game is to outscore your opponent, an increase in r_P or a decrease in r_A will have an identical impact on wins such that $f(r_P, r_A) = \delta_0 + \delta_1(r_P - r_A)$. These are the only two inputs necessary to produce wins. However, runs are scored and prevented at the team level; so both inputs must be further defined as functions of individual contributions in order to capture both productivity and behavioral effects. Specifically both r_{P_j} and r_{A_j} are defined as a function of three variables. First, total individual contributions (C_j) measures all inputs that can potentially be translated into output. Second, an index measuring the distribution of individual contributions, (h_j), captures the ease with which workers are able to complement individual inputs and collectively translate them into firm output. Lastly, an index measuring individual workers' behavioral response, (b_j), is an indicator of the average effort focused on the translation process, based on the working environment. This will capture influences such as salary disparity, that may hinder players' willingness to work together. Unlike (h_j) which captures cohesion of actual inputs, (b_j) measures a consciously chosen level of cooperation. This model adds to past literature by considering productivity effects and behavioral effects separately. The production process is fully contained within the baseball game. Once teams have been assembled, macroeconomics factors outside of the event have no bearing on the outcome. The talent of the players (C_j) and their ability and willingness to work together (h_j and b_j) should fully explain runs over the course of a season. This accounts for the level of offensive and defensive talent, relative to other teams in the league. The function is defined such that:

$$r_{P,j} = g(C_j, h_j, b_j) \tag{1.2}$$

and

$$r_{A,j} = k(C_j, h_j, b_j) \quad (1.3)$$

Lastly, I further identify C_j as a function of the realized production of each individual player, $c_{i,j}$ such that:

$$C_j = q(c_{1,j}, c_{2,j} \dots c_{n,j}) \quad (1.4)$$

Since the impact of an increase in r_P or a decrease in r_A is assumed to be identical, let $r_D = r_P - r_A$. This run differential can be interpreted as the individual contributions that the team can convert to wins, through what I call the translation process. From here I will test the hypothesis that r_D , and thus W_j , is increasing in c_i , but that h_j and b_j each have a unique impact on wins as well. The inclusion of h_j and b_j captures the distinction between player and team success. It is at this stage in the production process that individual contributions can be lost in translation. The impact of worker's behavioral response to salary disparity is borrowed from past literature. This model adds to the literature by separately considering the impact of the behavioral environment and productivity concentration, rather than using salary as a proxy for productivity.

1.4 Empirical Model, Variables, & Data

1.4.1 *Empirical Model*

In order to test the conceptual framework, I define the following functional form in which individual contributions have an isolated additive impact on

team output, $Wins_{j,t}$, such that:

$$Wins_{j,t} = \gamma(C_{j,t}, h_{j,t}, b_{j,t}) = \psi_0 + \psi_1 \left(\sum_{i=1}^n (c_{i,j,t}) \right) + \psi_2(h_{j,t}) + \psi_3(b_{j,t}) \quad (1.5)$$

The hypothesis is that $Wins_{j,t}$ is decreasing in both $(h_{j,t})$ and $(b_{j,t})$. The concentration of individual contributions affects team level output, independent of behavioral effects and the total level of individual contributions. In this cases, the more players contribute, the more opportunities the team has to convert those contributions to wins. However, as the concentration of contributions grows, or the working environment elicits less cooperation amongst players, less of the individual inputs can be converted into team level output, and is thus wasted.

To test the conceptual framework, I use an additive functional form rather than a multiplicative form, such as a Cobb-Douglas production function, because this provides more versatility. If individual contributions were multiplied together in the production function, disparity would have an implicit negative impact. An additive functional form separates and isolated the impact of concentration into the variable $(h_{j,t})$. This generalizes the form to allow concentration to affect different industries to varying degrees.

Single season wins is the dependent variable. Since teams appear each year, I organize the data into panels to capture its cross sectional time series nature. I record data for every MLB team from 1974-2014, but MLB has gone through three expansions during the span of the dataset, adding two teams in each 1977, 1993, and 1998. These teams do not appear in earlier years, so the panels are unbalanced. Groups are indexed by franchise and year. A new group is created when a team changes city, even if the franchise technically remains intact. For example, the Montreal Expos and the Washington Nationals are categorized as separate groups, despite being a single, continuous franchise. This is because,

when a team moves, I can reasonably expect the group to be characterized by a new set of unobservable, time-invariant, fixed effects, such as average regional climate, stadium details, substitutability of baseball entertainment, etc. The variable of interest is the concentration of Wins Above Replacement Player (WARP), a single metric used to calculate an individual player's contribution to team production. A wide array of potential contributions are considered when calculating WARP in order to represent the full scope of $c_{i,j}$.

One advantage to working with data on measurable contributions to the production process as opposed to using salary to measure productivity is that the direct productivity data are, in general, much less impacted by exogenous factors over time, such as inflation or popularity of the sport. Past literature has seen datasets that span a period of rapid growth in broadcasting revenue, wide swings in the team market value, and significant shifts in the bargaining power of the players union. These changes seriously impacted MLB salaries by inflating the value of MLB talent as well as restructuring the environment in which players could bargain for that value. Additionally, team effects can result in serious swings in how talent is compensated across teams. Revenue is much higher in, for example, New York City than in Tampa, though a home run does not do more good for a team on the East Coast than it does for a team on the West Coast or anywhere in between. Throughout the regular season, a run is equally important for all teams at any given time. However, it is still possible that systematic differences in team's ability to convert individual inputs to wins may linger between teams. By failing to control for these differences OLS most likely suffers from omitted variable bias. Unsurprisingly, a Breusch-Pagan Lagrangian Multiplier test returns a likelihood ratio test statistic of 7.11 and a corresponding p-value of .0038. From this, I can reject the null hypothesis that there are no systematic differences between teams. To begin testing my hypothesis, I first account for these differences by estimating a

fixed effects model. The model is shown in equation 1.6 below where μ_i represents team-specific effects:

$$\begin{aligned} Wins_{j,t} = & \phi_0 + \phi_1(WARPHHI_{j,t}) + \phi_2(TeamWARP_{j,t}) + \\ & + \phi_3(Strike1) + \phi_4(Strike2) + [X_{i,t}]\Phi_k + \mu_j + \epsilon_{j,t} \end{aligned} \quad (1.6)$$

$X_{i,t}$ is a vector of year fixed effects and team specific linear time trends. The standard errors are also bootstrapped and clustered at the team level.¹ When using MLB data, fixed effects are traditionally used because systematic differences among teams are intuitively time-invariant. These differences may include stadium, stability of climate and temperature, local cultural taste for the sport, and ownership style, population or media markets. However, the dataset used in this analysis covers four decades of game-play, which has the potential to capture long-run effects. As players get traded, sign with other teams, age, and eventually retire, clubhouse culture changes. These effects may linger for several years, but do not remain constant over time. The way that teammates work together can change dramatically over decades, as players and the coaching staff turn over.

One potential complication comes from the large number of players who are signed to multi-year contracts. It is reasonable to assume that the make-up of many teams will remain constant for short periods of time, but not over the entire dataset. Exceptional team chemistry can result in “dynasties”. In this situation management works to reduce turnover and maintain the winning combination of players and coaches for an extended period. Even when teams struggle, many of the players return the following year. This short-term stability of team make-up

¹The number of clusters bordered on “too few” based on Cameron, Gelbach, & Miller (2008). They recommend a wild cluster bootstrap to prevent over rejection. I use the Stata command `clustse` (Menger 2015) which performs the wild cluster bootstrap-t procedure as described by Cameron, Gelbach, & Miller. Statistical significance does not change using this method in any of the models outlined in the paper.

suggests this year’s level of production and concentration of both individual contributions and salary are all good predictors of next year’s production and concentration. If unobserved panel-level effects are correlated with the lagged dependent variable, then a fixed effects estimator will be inconsistent. Simply introducing an autoregressive term to the fixed effects model will likely result in dynamic panel bias (Nickell, 1981). Instead, in keeping with Kamada and Katayama (2014), I run an Arellano–Bond linear dynamic panel-data estimation, as shown in equation 1.7. Using generalized method of moments, this estimator identifies potential instruments for the lagged dependent variable in order to generate a more consistent estimator. For this model, I use a one step GMM process with robust standard errors.

$$\begin{aligned}
Wins_{i,t} = & \theta_0 + \theta_2(Wins_{i,t-1}) + \theta_2(WARPHHI_{i,t}) \\
& + \theta_3(TeamWARP_{i,t}) + \theta_4(Strike1) + \theta_5(Strike2) + [X_{i,t}]\Theta_k + \epsilon
\end{aligned}
\tag{1.7}$$

1.4.2 *Variables*

The dependent variable, $Wins_{j,t}$, is the total number of wins for team i in year t . This is equivalent to winning percentage, as the number of games scheduled does not vary by team or year, across the dataset.

Baseball is an attractive setting because of the detailed data on individual player contributions. As an example, consider the seminal work of Scully (1974), which uses Slugging Percentage (SLG) to evaluate an individual’s contribution in order to generate the first empirical analysis of marginal revenue product. Since Scully’s analysis, SLG has been a popular metric to calculate a player’s value, as it is easily calculated, yet strongly correlated with offensive team output. However, if one player is so productive that he single-handedly pulled up the team

SLG, it may be that this lone producer lacks the complements needed to convert his effort into runs. This would not be the case for another team with an identical team SLG, constructed entirely (or mostly) of players with roughly identical individual SLGs. As individual productivity becomes less concentrated, players can more easily find complements to their contributions and generate runs. Individual output that goes above and beyond the level with which co-workers can keep up becomes wasted effort. The strike out to base-on-balls ratio plays a similar role in the literature when considering the defensive contributions of pitchers. Both these measures are lacking in that they are both marked by an inability to view offensive and defensive contributions simultaneously. Wins Above Replacement Player (WARP) is calculated by *Baseball Prospectus* and is used as a catch-all to measure total individual player productivity. As such, WARP provides a more comprehensive view of overall performance, by considering the offensive and defensive contributions from pitchers and batters alike. One statistic that can combine offensive and defensive contributions allows for the combination of pitchers and position players in the same dataset. This is essential to gauge the concentration of talent and has a definitive advantage over statistics like SLG used in past literature.

WARP captures a wide array of contributions, including playing time, position, batting, base-running, and defense for batters, and role, innings pitched, and quality of performance for pitchers (*Baseball Prospectus* 2015). The team must collectively convert all of the individual contributions reflected in each player's WARP to team-level output, measured in wins. The idea of a "replacement player" can be thought of as a baseline level of talent. Most teams have top prospects who, although still in the minor league system, exhibit talent above replacement level. The replacement level, in theory, has to be low enough that it is constantly available in the minor leagues. Players with this level of talent should

have relative difficulty producing in the MLB environment, but are still capable of generating positive output. Such players are thus poised to “replace” -with the baseline level of productivity- any major league player who is injured, traded, or otherwise unable to perform. While WARP statistics are publicly available for all Major League players, the calculation details are proprietary. However, *Baseball Prospectus* does provide some insight into their calculations by discussing what contributions are considered and its own definition of replacement level. It estimates that a Major League team made up entirely of replacement players would win about 50 regular season games.

Before I can consider testing my hypothesis, I have to provide evidence that WARP can adequately represent the $c_{i,j}$ variable I defined in my conceptual framework. If *Baseball Prospectus*'s claims hold true, then regressing Wins on WARP should yield a constant of 50 and a coefficient of one on WARP. I test the model shown in equation 1.8, using pooled ordinary least squares. The model is run using teams' regular season wins from 1974-2014 as the independent variable and total team WARP as the only regressor. The claim that WARP is a complete catch-all for game-play performance should alleviate the need for additional controls.

$$Wins_{i,t} = \beta_0 + \beta_1(WARP_{i,t}) + \epsilon \quad (1.8)$$

The results in Table 1.1, show a constant of 48.14 and a coefficient of 1.07. While a t-test indicates that 1.07 is statistically different from 1, the definition is, for all intents and purposes, accurate. The same holds for the constant.

WARP also makes sense intuitively. The expectation that a zero WARP team would win about 1/3 of its games is reasonable. This would be one of the worst regular season records in the modern professional baseball era. However,

Table 1.1: OLS Regression of $W_{i,t}$ on WARP

VARIABLES	(1) w
$WARP_{i,t}$	1.073*** (0.0220)
Constant	48.14*** (0.676)
Observations	1,138
R-squared	0.676
Standard errors in parentheses	
*** p<0.01, ** p<0.05, * p<0.1	

players so close to MLB caliber should be expected to perform with at least some degree of success. A player's WARP indicates how much value, beyond replacement level, he adds to his team. If the addition of a new player increases the sum total of a team's WARP by one, that team's expected wins should also increase by one. A particularly poor performance by an MLB player over the course of a season could result in a negative WARP. A negative value indicates that the team would have been better off using a replacement-level player instead.

Further, a correlation table allows me to compare the statistic to other commonly used productivity measures. Table 1.2 shows that, at the team level, total WARP has a much higher correlation with wins than other standard measures of performance including batting average (AVG), on-base percentage (OBP), slugging percentage (SLG), walks and hits per inning pitched (WHIP), and earned run average (ERA). This is unsurprising, as all other variables measure either offensive or defensive performance. Since WARP is constructed using components from a much wider variety of metrics, it is able to capture more of the variation in wins. This makes it a more attractive choice as the individual input variable, $c_{i,j}$. It is highly correlated with wins, but derived from individual inputs, rather than collective team output. This analysis is pooled and views all of MLB across

the entire dataset. A closer look at concentration as well as team-level effects and year fixed effects will yield more accurate estimates of what drives wins.

Table 1.2: Cross-Correlation Table - Team Data 1974-2014

Variables	W	WARP	AVG	OBP	SLG	WHIP	ERA
W	1.000						
WARP	0.822	1.000					
AVG	0.362	0.558	1.000				
OBP	0.441	0.652	0.846	1.000			
SLG	0.387	0.611	0.752	0.762	1.000		
WHIP	-0.535	-0.377	0.213	0.204	0.254	1.000	
ERA	-0.456	-0.270	0.289	0.285	0.434	0.897	1.000

I gather data on each team's top 15 WARP contributors from 1974-2014. I choose a cutoff of 15 players on each team because I can reasonably expect this to be enough players to represent a team's core producers. 15 is also a natural break in the data, where negative WARPs just begin to appear. Frequently, these core producers include: 8 position players, 5 starting pitchers and one closer as well as a 15th spot for a middle reliever or a designated hitter. Alternatively, a team with a philosophy that favors defense, may invest in top tier relief pitching, such that multiple pitchers squeeze position players out of the top 15. Regardless, the inclusion of 15 players reasonably covers the significant contributors to a team.

The concentration is measured by a Herfindahl-Hirschman Index which I call WARPHHI. This is a popular measure of concentration for its simplicity in calculation and intuitive interpretation. All HHI results are bound between 0 and 1, with 1 representing total concentration. 0 represents perfectly even distribution, though is only achieved in the limit as the number of firms approaches infinity. The HHI is calculated by first taking each player's WARP and dividing it by the total WARP of his team. For each team, the ratios are then squared and summed. When larger shares are held in fewer hands, the sum of squares grows.

Negative numbers cannot be used to calculate HHI, since percentages are squared and a player who posts a WARP of -1 would be indistinguishable from a player who posts a WARP of 1. Additionally, if negative WARP were to detract from the sum total of contribution, it would be possible for a player to contribute more than 100% of individual production. To avoid these issues, negative WARP is censored at zero, and each player is assigned a share of the sum total of positive WARP values. The cutoff of 15 leaves only five players, of 17,070, with negative WARP that must be censored at zero. Of these five players, three are from the 1985 Braves and none of the WARPs drops below -0.42. As only three of 1138 teams are impacted, this censoring is likely to be minimally disruptive.

I start from 1974 because this was the first year of MLB's current salary arbitration process, which changed the way that teams allocate payroll, and potentially forced a change in the distribution of talent across the league. Prior to 1974, the reserve clause bound players to the team that drafted them. Each year, teams take turn drafting players. The league forces teams to respect all other teams' draft picks, and once a player is drafted, he is unable to entertain offers from another team. If he refuses to sign with the drafting team, he must wait until the following year's draft. The reserve clause allowed teams to reserve the right to renew any contract for one year. Seemingly benign, this actually allowed teams to renew every aspect of the contract, including the right to renew; thereby allowing teams to renew player contracts in perpetuity. This gave teams complete control over player mobility and total monopsony power. This system severely suppressed wages for players. The creation of the arbitration system in 1974 greatly reduced the power of the reserve clause, by providing a time-line for a shift in bargaining power. The reserve clause became limited to three years. After this time period, if a player and team could not agree upon a contract, either the team or player can defer to an independent panel of arbitrators to determine binding salary terms for

the next season. Two years later, players gained even greater bargaining power by negotiating a system that allowed players to earn free agency. In the current system, after three years of arbitration eligibility, players reach free agency, in which they can entertain offers from any team in the league. The birth of this system restructured the productivity and behavioral effects relevant to this analysis.

$WARPHHI_{i,t}$ measures the concentration of productivity as measured by WARP for team i in year t . $TeamWARP_{i,t}$ is the total cumulative WARP for team i in year t , which represents the sum of individual contributions that can be converted to team-level output, measured in wins.

SalaryHHI captures the behavioral response to salary concentration and is calculated in a fashion similar to WARPHHI. The top 15 salaries are considered. Due to limited salary data, this reduces the years available to 1985-2004.

Controls include total WARP for the 15 highest ranking players on the team ($TeamWARP$), total salary for the 15 highest paid players ($TotalSalary$), team specific linear time trends and year fixed effects. Of particular note are the years 1994 and 1995 because of a player's strike that cut the '94 season short and delayed the start of the '95 season. These particular years are denoted $Strike94$ and $Strike95$ and are reported in subsequent tables.

Table 1.3: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Wins	79.483	12.387	37	116	1138
TeamWARP	30.761	7.264	7.21	52.39	1138
WARPHHI	0.1	0.018	0.069	0.225	1138
TotalSalary*	33.96	24.06	3.73	165.75	558
SalaryHHI	0.103	0.028	0.069	0.269	558

*in millions

1.4.3 *Data*

Performance data, including wins and player productivity, were gathered from *Baseball Prospectus* (2015). Salary data comes from the Society of American Baseball Research (Pappas, 2004). The complete dataset is an unbalanced panel of 32 groups over 40 years for a total of 1138 observations.

1.5 Results

Table 1.4 presents results from regressions using Ordinary Least Squares fixed effects, and Arellano-Bond estimators respectively. WARPHHI has a negative coefficient in each of the models, indicating that lower individual productivity concentration is associated with better team performance, holding overall talent constant.

Once I control for WARPHHI, both strike years, and team effects, the coefficient on TeamWARP increases in all models compared to the OLS model in which TeamWARP was the only regressor. As expected, Strike1 and Strike2 both have negative coefficients (other than Strike95 in the fixed effects model, where it is insignificant), as fewer games were played during these years. TeamWARP and WARPHHI, are significant at the 99% level. Lagged wins in the Arellano-Bond model, is significant at the 90% level. Though I use year fixed effects, only the coefficients for the years 1994 and 1995 are shown (labeled Strike1 and Strike2), as a major negative impact was expected.

Table 1.4: Estimation with Direct Productivity Measures

Dependent Variable is single season Wins, W_t			
	(1)	(2)	(3)
VARIABLES	OLS	Fixed Effects	Arellano-Bond
W_{t-1}			0.0660*** (0.0249)
TeamWARP	1.264*** (0.0298)	1.249*** (0.0379)	1.232*** (0.0382)
WARPHHI	-44.75*** (11.94)	-49.55*** (11.85)	-49.65*** (11.34)
Strike94	-17.08*** (2.080)	-5.180 (5.563)	-17.20*** (1.196)
Strike95	-8.036*** (2.232)	4.508 (5.868)	-6.397*** (1.355)
Constant	46.94*** (2.205)	47.12*** (2.553)	43.60*** (2.483)
Year Fixed Effects	yes	yes	yes
Team Specific			
Linear Time Trends	yes	yes	yes
Observations	1,138	1,138	1,074
R-squared	0.754	0.745	
Number of Teams		32	32

Bootstrapped Standard errors in parentheses. Clustered at the team level .

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The analysis simply replaces behavioral effects with productivity effects, rather than separates them. To capture both effects, I repeat the above estimation with the addition of TotalSalary and SalaryHHI included as regressors to more accurately reflect the theoretical model. Before proceeding, I must consider that salary and WARP may be correlated. The variance inflation factor is used to determine the extent of any multicollinearity. Results do not indicate that multicollinearity should be a concern and are shown in Table 1.5.

Table 1.5: Variance Inflation Factor

VARIABLE	(1) VIF	(2) 1/VIF
TeamWARP	1.22	0.820173
SalaryHHI	1.08	0.921956
Total Salary	1.08	0.927817
Strike1	1.06	0.945677
Strike2	1.04	0.965846
TeamHHI	1.02	0.978452
Mean VIF	1.08	

Interestingly, TotalSalary is not significant in any of the models in Table 1.6. Salary, when used as a proxy for productivity, has a meaningful and statistically significant affect on wins. This is the norm in past literature, and similar results are recreated at the end of this section. Here however, once I control for WARP, total salary does not appear to have any impact. Salary concentration, on the other hand, does appear significant in all models, even when controlling for total WARP and WARP concentration. This indicates that any relationship between wins and total salary reflects salary's weak correlation with WARP; but salary concentration reflects behavioral responses independent of productivity. This supports the assertion that productivity concentration and salary concentration should be viewed separately. Full results are shown in Table 1.6.

TeamWARP is positive and significant at the 99% level in all models. TeamHHI and SalaryHHI are both negative and significant at, at least, the 95% level in all models. As expected, both dummy variables for years of the players' strike are negative in each model and significant in each case other than Strike95 in the fixed effects model. The respective coefficients on all dependent variables are of approximately the same magnitude across the three models.

Table 1.6: Estimation with Direct Productivity and Salary Measures
 Dependent Variable is single season Wins, W_t

	(1)	(2)	(3)
VARIABLES	OLS	Fixed Effects	Arellano-Bond
W_{t-1}			0.0257 (0.0358)
TeamWARP	1.243*** (0.0464)	1.274*** (0.0555)	1.243*** (0.0544)
WARPHHI	-50.96*** (17.06)	-42.20** (18.23)	-43.89** (19.68)
TotalSalary [†]	0.0244 (0.0263)	-0.0144 (0.0299)	-0.0451 (0.0336)
SalaryHHI	-30.98** (12.93)	-37.03** (14.97)	-32.34** (13.91)
Strike94	-17.16*** (2.107)	-14.74*** (3.262)	-15.80*** (1.638)
Strike95	-7.486*** (2.346)	-4.801 (3.601)	-5.301*** (1.758)
Constant	51.49*** (3.251)	53.34*** (4.365)	51.67*** (5.807)
Year Fixed Effects	yes	yes	yes
Team Specific			
Linear Time Trends	yes	yes	yes
Observations	558	558	524
R-squared	0.775	0.762	
Number of Teams		30	30

[†] in millions

Bootstrapped Standard errors in parentheses. Clustered at the team level .

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

These results can be contrasted with a more traditional approach that does not control for productivity directly. The methods employed to obtain the following results are identical to those discussed previously, save for the omission of TeamWARP and WARPHHI with results shown in Table 1.7. When compared to the full model, these results show a large increase in the magnitude of the coefficient on SalaryHHI_t. TotalSalary_t remains insignificant in the fixed effects

and Arellano-Bond models. Lagged wins is significant at the 95% level. These results confirm the importance of including productivity effects in the model before drawing any conclusions about behavioral effects. As expected, omitted variable bias seems to be largely inflating the coefficient on SalaryHHI in all models.

Table 1.7: Estimation without Direct Productivity Measures

Dependent Variable is single season Wins, W_t			
	(1)	(2)	(3)
VARIABLES	OLS	Fixed Effects	Arellano-Bond
W_{t-1}			0.160** (0.0683)
TotalSalary [†]	0.134*** (0.0459)	0.0691 (0.0592)	-0.0445 (0.0633)
SalaryHHI	-78.38*** (22.35)	-84.64*** (22.95)	-71.98*** (18.25)
Strike94	-23.71*** (3.213)	-29.36*** (6.822)	-24.38*** (2.166)
Strike95	-7.371* (3.767)	-13.71* (7.748)	-4.640 (2.893)
Constant	87.27*** (3.473)	79.30*** (8.391)	72.88*** (8.191)
Year Fixed Effects	yes	yes	yes
Team Specific			
Linear Time Trends	yes	yes	yes
Observations	558	558	524
R-squared	0.409	0.405	
Number of Teams		30	30

[†] in millions

Bootstrapped Standard errors in parentheses. Clustered at the team level .

*** p<0.01, ** p<0.05, * p<0.1

Results indicate that the range of concentration in WARP and Salary can explain over six wins each, when controlling for total WARP and total salary. In a league in which playoff berths can be decided by a single win, this represents a very large swing. Six wins can be the difference between a top seeded playoff contender and missing the playoffs entirely. Appendix A provides lorenz curves as a visualization of the extent to which teams WARP concentration may vary. Appendix B provides a similar analysis using gini coefficients instead of HHI as a robustness check.

1.6 Discussion & Conclusion

As firms seek to efficiently translate individual contributions to firm-level output, they must consider the concentration of individual labor productivity. MLB data show that an increase in the concentration of individual production has a significant negative effect on collective team performance, independent of behavioral effects. The distinction between production concentration and salary concentration is an important one. The talent of a superstar is likely wasted when he is asked to carry a team of relatively weak teammates. The extent to which individual inputs complement one another will determine the degree to which productivity concentration affects firm performance. A strategy of incrementally rounding out talent is superior to building a team around a limited selection of stand-out contributors. When inputs are complementary, as appears to be the case in MLB, the production function should indicate a negative relationship between output and the concentration of individual contributions. When hiring workers, firms must consider not only the expected productivity of the worker, but also how the addition to the labor force will change concentration. How the distribution

of talent will impact collective productivity will help teams determine how much they are willing to offer free agents.

Baseball has a very rigid production process. For example, each player must await his turn to bat. Rules prohibit a team from having its best hitter step up to the plate more often. The analysis may extend to a sport like basketball in that a star player must rely on teammates to complement his talent, move the ball, screen defenders, get rebounds, provide outlets for assists, etc. However, unlike baseball, a basketball team can plan its game strategy such that its most talented offensive player takes more shots. Similarly, a football team has the ability to shift focus toward a running or passing game depending on its distribution of talent. In these cases, talent may exhibit elements of both complements and substitutes. This will cause the affect of concentration to vary in magnitude. Despite these specific environments, the necessity to consider talent distribution should apply to the general case.

This strategy of considering potential productivity can expand to any industry in which co-workers perform complementary tasks to achieve a production goal that is measured as a collective result. For example, each floor of a hospital must decide on the mix of doctors, nurse practitioners, and registered nurses. Determining the productivity distribution that optimizes patient recovery rates will aid in determining the most efficient allocation of the wage bill. This contrasts with past literature that suggests a reverse approach in which that the wage bill can be used to manipulate productivity. Firms should not attempt organize the wage bill in order to optimize productivity per worker, nor should they attempt to manipulate the distribution of talent be organized to optimize output per dollar. Results here indicate that neither approach is comprehensive enough. Productivity effects and behavioral effects play unique roles in the production function, and must be considered separately.

1.7 Future Work

The use of HHI in this analysis may spark criticism, as the index fails to provide information on where the concentration is focused. As an example, consider a team for which a single player contributes the bulk of the team's productivity as well as the bulk of the payroll. Compare this to a team for which a single player contributes the bulk of productivity, but a different player claims the bulk of the payroll. Both of these teams would appear with two high HHIs and be indistinguishable from one another in the data, despite offering very different working environments. This is the unfortunate nature of the Herfindahl-Hirschman index. This paper has shown the importance of separating productivity and behavioral effects. A potential extension would be to further define the behavioral environment by developing a metric that considers salary relative to WARP. It may be that some salary disparity is viewed as equitable, provided the players who earn larger salaries justify their pay with added productivity. Further extensions would explore the applicability of this analysis to additional industries. Many industries provide additional challenges to the analysis, as they typically have less data on individual performance. Still, instruments such as education, years of experience, piece rates, or various specific individual success rates may allow for the separation of performance from salary when calculating concentration. Unobservable effects typically plague the correlation between wage and performance. Understanding how worker contributions interact separate from wage allows firms to draw a more direct link between individual performance and team performance. This will, in turn, lead to a cleaner decision making process when it comes to allocating the wage bill.

CHAPTER 2

FIRM BEHAVIOR UNDER CONDITIONS OF FINAL OFFER ARBITRATION: AN EMPIRICAL ANALYSIS OF MAJOR LEAGUE BASEBALL'S SALARY NEGOTIATION STRUCTURE

2.1 Introduction

Interest arbitration, the practice of deferring to a third party to resolve labor disputes that arise during private negotiation, is intended to prevent work stoppages, as work stoppage is an inefficient bargaining tool. Reaching terms after a stoppage is inferior to reaching identical terms prior to a work stoppage. If the negotiating parties cannot settle their dispute privately by a set deadline, they must defer to an arbitrator to resolve the impasse. This independent party assigns terms for the next labor contract. Arbitration has been particularly useful in the public sector for industries in which a work stoppage presents a public safety hazard, as is the case for police officers and firefighters. It is also helpful for negotiations involving medical personnel when a work stoppage risks harm to patients.

Farber and Katz (1979) cite two criteria by which one can evaluate arbitration: infrequency of use and resemblance to privately negotiated contracts. Further, they analyze interest arbitration and its impact on the incentive to negotiate. They assert that the use of arbitration as a negotiation tool results from uncertainty surrounding the arbitrator's decision. In this way, the arbitration procedure defines the contract zone. Each party forms expectations about the arbitrator's notion of a fair settlement. Parties that are more risk averse will be willing to privately negotiate a settlement that is less than the expected settlement from interest arbitration. Therefore, when faced with interest arbitration, a negotiated settlement favors the less risk averse party. When neither party can perfectly forecast an arbitrator's ruling, the arbitrator's ability to promote private bargaining increases.

In order to curtail any bias that may stem from an arbitrator's personal preferences, a selection mechanism gives both parties equal weight in choosing the arbitrator who will be hired to hear the case, thereby incentivizing arbitrators to find an equitable settlement. Any arbitrator perceived to systematically favor one party or the other will likely be vetoed by the disadvantaged party in all future proceedings. If an arbitrator seeks future work, then he or she cannot allow rulings to reveal any forecastable bias. The losing party will see any bargaining power that stems solely from the arbitrator as undue. The incentive to be attractive to both parties in order to secure future work forces arbitrators to consider the most mutually acceptable outcome.

This paper analyzes binding final offer arbitration (FOA), a specific form of interest arbitration introduced formally by Stevens (1966). FOA differs from conventional arbitration (CA), in which the arbitrator is free to formulate a compromise at his discretion, which may or may not be binding.² When parties

²Non-binding arbitration uses a similar process, but the arbitrator's role there is to provide

agree to operate in a negotiation environment where FOA is the conflict resolution mechanism, then if negotiating parties are unable to reach mutually satisfactory terms by a set deadline, firms and workers bring their respective terms to the hearing, and the arbitrator must side with one party or the other, writing that party's terms into a binding contract without room for compromise. The uncertainty surrounding the arbitrator's decision creates an incentive for parties to privately negotiate. Further, when a hearing is necessary, outcomes should fall in the same range as those expected from private negotiation. The direct cost associated with hiring an arbitrator, as well as the indirect cost of the positive probability that the arbitrator will side with an opponent should incentivize private bargaining. If an impasse is still reached, the FOA mechanism resolves the dispute without a work-stoppage.

One major criticism of CA, that FOA was designed to address, is that arbitrators attempt to secure future work by refusing to take sides and simply "split the difference" between the offers brought to the arbitration table (Farber 1981). Instead of determining which party has prepared more justifiable and reasonable terms, and providing that party with the added weight that their argument deserves, "splitting the difference" strikes a compromise that invariably gives equal consideration to both sides, regardless of evidence or justification. With a clear notion of how the arbitrator will act, the parties no longer have an incentive to negotiate. Instead, they have an incentive to bring extreme offers to the arbitrator in hopes of tipping the average in their favor. This causes two problems. First, it chills bargaining. Parties focus on influencing the arbitrator instead of negotiating with one another. Second, it reduces the usefulness of the information the arbitrator receives in the hearing. The extreme arguments brought to the table in an effort to shift the average do little to inform the arbitrator of the worker's

information for future negotiation, and the arbitrator's decision does not necessarily dictate the settlement.

true value. In this situation, the arbitrator has become the focus of both parties' negotiation strategy, yet as a third party, he receives poor information from which he must glean the worker's value.

If FOA is employed instead of CA, then rather than trying to sway the average, firms and workers are locked in a competition to see who can appear more reasonable to the arbitrator. This creates an added incentive to negotiate privately. Parties know that losing an arbitration hearing amounts to totally conceding to the other party's terms. Secondly, each party's offer will more closely reflect fair market value. This gives the arbitrator more reliable information on which to base his or her decision.

This paper uses data from Major League Baseball's binding FOA framework to explore whether arbitrated settlements resemble privately negotiated settlements. Theory suggests that neither party should expect arbitration to yield a settlement outside of the expected range of privately negotiated salaries. However, results in this paper indicate that avoiding an arbitration hearing allows players to leverage future bargaining power to extract salary gains.

The rest of the paper proceeds as follows: Section 2.2 provides additional background on my research question, motivation, and why MLB data is particularly useful for this analysis. Section 2.3 discusses related literature. Section 2.4 outlines a brief conceptual framework that will be tested in the empirical model. Section 2.5 introduces data and variables. Section 2.6 includes the empirical model and results. Section 2.7 discusses and concludes the paper.

2.2 Institutional Background

Major League Baseball (MLB) provides a framework in which FOA is available for a short period during a player's career. Prior to arbitration eligibility, teams own exclusive rights to a player's contract. As expected, teams can use this monopsony power to suppress salaries. The right to FOA is typically earned after three years of Major League service and is the first step in breaking down team monopsony power. After six years, FOA is replaced by free agency, a competitive setting in which players can more freely sell their services³. This tiered system of full monopsony power, FOA, and free agency allows me to compare management's and labor's behavior when utilizing FOA and their behavior in an alternative private negotiation environment.

MLB's use of FOA differs from public sector use in that it is invoked when teams are negotiating with an individual player, not the entire union. In addition, when either party calls for a hearing, MLB guidelines restrict proposed compensation packages during arbitration hearing to a single season salary. Arbitrated terms cannot include multi-year contracts, performance bonuses, or any other alternative compensation. An FOA system should force parties to develop arguments that accurately reflect the market value of the player.

Additionally, teams and players often provide public information on the terms they intend to present during an arbitration hearing, even if they settle before the hearing takes place. In these cases, I can use data that include both the arbitration strategy and private negotiation results for a single case and analyze how they might differ. This paper uses data from MLB to test the hypothesis that FOA settlements are statistically identical to privately negotiated settlements.

³See MLB Collective Bargaining Agreement 2012-2016, available from The Major League Baseball Players Association (<http://www.mlbplayers.com>).

FOA's ability to mimic privately negotiated settlements without a work stoppage is vital to its efficacy as an impasse resolution mechanism.

Historically, baseball's exemption from anti-trust laws, consistently and inexplicably upheld by the Supreme Court, allowed MLB to act as a cartel. Teams exercised full monopsony power over the players they drafted, as all other teams in the league would honor each other's contracts⁴ The contracts included a "reserve clause," which tied players to their teams in perpetuity at the teams' discretion. Players had no path to free agency, nor was there a legal avenue to sue the league on anti-trust grounds. Eventually, the Major League Baseball Players Association (MLBPA), gained enough power as a union to chip away at the teams' monopsony power.

The MLB Players' Association (MLBPA) successfully added FOA to the collective bargaining agreement in 1970. The first round of eligibility went into effect in 1974. Two years later, the union negotiated terms that allowed players to earn full free agency. Two players, Andy Messersmith and Dave McNally, won the right to free agency when they filed grievances against MLB, arguing that the the reserve clause should represent the team's right to renew a contract for a single year, not in perpetuity. This limit on the reserve clause would provide a clear path to free agency. The grievance was heard by a panel of 3 arbitrators: Marvin Miller representing the union, John Gaherin representing the owners, and Peter Seitz chosen by both parties as an impartial arbitrator. After Seitz ruled in favor of the players, the owners immediately exercised their right to fire him (Miller 1991).

⁴This precedent for MLB's antitrust exemption was established following the dismantling of the Federal League after only two seasons (1914 & 1915). Organized baseball was sued on antitrust grounds by Ned Hanlon, a Federal League owner from Baltimore. The case reached the Supreme Court which ruled that baseball did not operate as interstate commerce, and thus was not subject to antitrust laws. See Abrams (1998)

Currently, when a MLB player is drafted, he is tied to the drafting team, unable to entertain offers from others, but the team's monopsony power is temporary. Most players typically develop in the minor leagues for a number of years before being called up to play at the major league level. During the player's time in the minor leagues, and for his first three years of Major League play, the team retains full monopsony control over him⁵.

For a Major League player's fourth through sixth years of service, he is still tied to his team but is now eligible to renegotiate his contract under more favorable circumstances. If mutually satisfactory terms cannot be reached privately, either party may call for a FOA hearing.⁶ The team and player each present a salary offer for the following season along with their respective cases in support of their proposals. The arbitrator then selects one of the two offers, with no option for a compromise. The salary is binding for the coming season, and the contract is renegotiated the following year. During negotiations prior to the arbitration hearing, teams are free to offer multi-year deals, performance-based incentives, bonuses, and a variety of alternative pay structures, with arbitration exercised only at an impasse.

After a player has served 6 years, he becomes a free agent and can sell his services to the highest bidder. This breaks the monopsony of one team into

⁵An exception to this is MLB's Rule 5 draft. Players, not on an MLB team's 40 man roster, who were signed when they were 19 or older and have played in professional baseball for four years are eligible, as are players who were signed at 18 and have played for five years (mlb.com)

⁶In some cases, a player may become arbitration eligible after only 2 years. This is known as MLB's "Super Two" rule. According to Article VI(E)(1) of the MLB collective bargaining agreement: "Super Two" Players. In addition, a Player with at least two but less than three years of Major League service shall be eligible for salary arbitration if: (a) he has accumulated at least 86 days of service during the immediately preceding season; and (b) he ranks in the top 22% (rounded to the nearest whole number) in total service in the class of Players who have at least two but less than three years of Major League service, however accumulated, but with at least 86 days of service accumulated during the immediately preceding season. If two or more Players are tied at 22%, all such Players shall be eligible.

an auction involving 30 competing teams. Once a player reaches free agency, the market should begin to resemble a perfectly competitive market, with players begin paid close to their MRP. The presence of uncertainty, especially since salaries are determined prior to performance, may generate salaries greater than, or less than, MRP. Overpayment comes in two forms. First, injury or an unexpected decrease in ability will result in overpayment. Secondly, the free agent market resembles an auction, as parties bid against one another for talent. This can give way to the winner's curse, as teams overbid in an effort to secure players (Cassing, Douglas (1980), Thaler (2012)). Underpayment results from expectedly positive performance. A bargaining structure that includes arbitration as MLB's does, provides useful insight into negotiation behavior when parties use FOA as their impasse resolution mechanism. Most notably, players and teams often announce their official FOA arguments well in advance of a hearing. A large majority of these contracts are settled privately after the figures are announced but before an FOA hearing is necessary. This revelation of FOA arguments even when a private settlements occurs allows for interesting insight into how parties' FOA strategies differ from privately negotiations.

2.3 Related Literature

To develop a model, I draw from theoretical and empirical literature concerning party strategy and arbitrator behavior when faced with a negotiation environment in which arbitration is the agreed upon conflict resolution mechanism. Farber (1981) details two tactics that an arbitrator may employ when settling a dispute, Equitable Compromise and Equitable Settlement. In both strategies, the arbitrator weighs the firm's terms (Offer), the worker's terms (Ask), and his or her own understanding of a fair settlement. The major implication of Farber's theoretical analysis is that "splitting the difference" is the Nash Equilibrium of

this bargaining game. Parties arrange their terms around their expectations of the arbitrator's valuation. This makes it appear that arbitrators are compromising when, in fact, causality goes in the opposite direction. The parties respond to the arbitrator's notion of a fair settlement, instead of the arbitrator's responding to the parties' terms. Bazerman and Faber (1985) use simulations to identify how arbitrators consider the facts of a case as well as the offers presented in order to formulate settlements. This sheds light on how parties might use their terms strategically when trying to influence the outcome of an arbitrated settlement. Faurot and McAllister (1992) and Miller (2000) elaborate on this idea of how parties can use their proposed terms to influence outcomes specifically in the context of MLB's FOA negotiation environment. Since players and teams exchange figures well in advance of an arbitration hearing, it can be shown that parties likely use the announcement of terms as a means to maximize the expected returns from the negotiation that occurs in-between the announcement and the hearing. In this environment, exchanged figures provide vital information about the negotiation process, even if a hearing is avoided.

Bloom and Cavanagh (1984) and Ashenfelder and Bloom (1984) both offer results suggesting a tendency for both parties to adopt a more conservative stance when they use FOA instead of CA. Coleman, Jennings and McLaughlin (1993) draw the same conclusions when estimating a model using MLB data by showing that more conciliatory offers do little to impact the expected monetary value of the settlement. This provides an incentive when facing FOA for higher initial offers from firms and lower requests by workers. Ashenfelder and Bloom (1984) goes on to study union negotiations in New Jersey, where unions win far more often in FOA. They conclude that this imbalance in wins is not due to arbitrator bias, but unions' systematically taking a more conservative approach than firms to formulating their terms. This stems from asymmetric risk aversion.

Despite FOA's effectiveness in promoting private negotiation, Burgess and Marburger (1993) find that FOA cases won by management typically yield salaries that are 9% lower than for comparable players who negotiated their salaries privately. The difference in outcome is statistically significant when controlling for differences between pitchers and position players, denoted hitters. Similarly, FOA cases that are won by players are 14% higher for pitchers and 4% higher for hitters than for comparable players who negotiated their salaries privately, though the increase for hitters is not statistically significant. This contrasts with previous empirical findings that FOA leads to more conciliatory offers. Babcock and Taylor (1996) support these findings by making use of an experiment conducted by Del'Omo (1987, 1989) in which each of 22 arbitrators is given the same 32 hypothetical cases and asked how he or she would rule. They generate a measure of unpredictability by comparing each arbitrator's decision to the average ruling for an identical scenario. Results show that an increase in uncertainty reduces the use of arbitration, however unpredictable arbitrators may produce low-quality settlements in the event that a private negotiation impasse arises. Here, "low-quality" means that an arbitrator may unexpectedly side with an offer outside the range in which a privately-negotiated settlement would lie. These findings do not necessarily extend to conventional arbitration as shown by higher quality arbitrated settlements observed in the National Hockey League (Lambrinos, Ashman 2007).

While it is important to note how the behavior of the negotiating parties change when faced with the option of arbitration, one must also consider the behavior of the arbitrators themselves. Farber and Bazerman (1986) introduce a model in which an arbitrator's decision is a function of the terms presented by each party during the hearing, as well as the arbitrator's personal determination of an appropriate award based on his understanding of the dispute prior to the hearing.

However, the arbitrator also wants to maximize the probability of being rehired in the future. Wittman (1986) provides a theoretical framework for determining a party's strategy when facing a scenario in which an arbitrator allows bias to shift his preferred award toward one party. He also considers how parties' strategies change when an arbitrator alters his relative sensitivity to offers. Marburger and Burgess (2004) and Marburger (2004) extend these theoretical models with an empirical analysis of how arbitrators determined preferred awards. Empirical findings are of particular interest because they provide evidence that arbitration eligibility creates a unique distribution of settlements that falls below the distribution of settlements in the free agent market, but above the purely monopsonistic market.

Lastly, the asymmetric risk aversion mentioned in Ashenfelder and Bloom (1984) plays an important role in motivating the analysis modeled in the next section. It is common to assume workers are more risk averse than firms. Frederick, Kaempfer, Ross Wobbekind (1998) use MLB data to explore why players chose to arbitrate and how avoiding a hearing can mitigate risk. Alternatively, Maxcy (2004a) theorizes that long-term contracts may be a way for firms, not workers, to mitigate risk depending on the relative uncertainty of market value and productivity. Further, MLB data support this theory (Maxcy 2004b). MLB provides a system in which arbitrated cases may only involve single-season salary. A multi-year contract signed by an arbitration eligible player necessarily represents an avoidance of an arbitration hearing. Additionally, a multi-year contract also likely involves the purchases of year, or years, in which the player would have otherwise been free to sell his services in the free agent market. This allows for an interesting analysis of negotiation behavior.

Building on this literature, the next section constructs a model in which the player's *Ask* and the team's *Offer* play an important role in the negotiation

process, even if a hearing is avoided. The length of the contract provides useful information on relative risk aversion, identifying who pays a premium for contract extensions. I add to the literature by exploring new, direct measures of productivity. Additionally, I estimate the model using both an OLS and fixed effects estimator, and then argue that the environment is more appropriately estimated with a quantile regression analysis.

2.4 Conceptual Framework

I design a framework that outlines negotiation between a worker and a firm when FOA is used as a impasse resolution mechanism. From this model, I construct an empirical framework to test the hypothesis that negotiations in the presence of arbitration yield settlements that resemble privately negotiated settlements. Since my empirical test uses MLB data, I use “team” to designate the firm, and “player” to designate the worker. Let $A_{i,j}$ denote the negotiated result for the average yearly salary of player i on team j . $A_{i,j}$ is defined as a weighted average of player i 's target salary, W_i , and team j 's target offer, F_j , such that:

$$A_{i,j} = f(F_j, W_i) = \delta_1(F_j) + \delta_2(W_i) \quad (2.1)$$

$$0 < \delta_1 \leq 1,$$

$$0 < \delta_2 \leq 1$$

and

$$\delta_1 + \delta_2 = 1$$

The result $A_{i,j} = F_j$ represents complete acquiescence by the player to the team's terms. Similarly, $A_{i,j} = W_i$ represents total concession by the team. Thus the distance from F_j to W_i represents the contract zone, and the deltas determine the exact resolution.

F_j is a function of team j 's expectation of the value of player i 's future performance, U_i , and the length of the contract, L , and time, t . The team's target wage is the absolute low end of the contract zone, such that its acceptance by the player indicates that the team has a very strong bargaining position. The team is always willing to accept an "Ask" less than F_j , though I assume that information is complete enough that a player would not inadvertently suggest a W_i less than the firm's F_j . The team is also willing to pay more than F_j , except in cases of the most extreme team bargaining power ($\delta_1 = 1$).

$$F_j = h(U_i, L, t) \tag{2.2}$$

I define W_i as a function of player i 's forecast of his own performance in the next season, denoted V_i , and the length of the contract, L , and time, t . Analogous to the team, the player's target wage is the absolute top of the contract zone, such that its acceptance by the team indicates that the player has a very strong bargaining position. The player will always accept more than W_i , but the team is unlikely to inadvertently reveal an F_j greater than the player's target salary. Except in cases of extremely strong player bargaining power ($\delta_2 = 1$), the player will also accept less than W_i .

$$W_i = g(V_i, L, t) \tag{2.3}$$

Since contracts are signed before players produce, decreased performance due to injury or wear on the body represents a risk for both player and team. Poor performance by a player represents an immediate loss of value for the firm, as well as a decrease in bargaining power for the player in future negotiations. Bearing these risks in mind, the parties must consider the total yearly salary established in the contract as well as the length of the contract. Teams and players must also consider a positive probability that an injury or wear on the body reduces a player's value in the future.

If players are more risk averse than teams, then when negotiating multi-year contracts, the target average yearly salary is decreasing in L for both players and teams. Teams discount future productivity. Since teams are willing to offer less for marginal years, the average yearly salary falls even as the total value of the contract grows. So long as contracts are guaranteed, as is the case for MLB, players are willing to trade average yearly salary in exchange for a larger total. For example, consider a situation in which a player and team both expect a player's performance to be worth \$1M in each of the next two seasons with a 10 percent chance in each season that the player will suffer an injury that reduces all future value to zero. If both parties are risk neutral and have perfect access to this information, then they will be indifferent between signing \$1M contracts one season at a time, and signing a 2-year contract worth a guaranteed \$1.9M. The second figure comes from taking the \$1M guaranteed for year one and adding \$1M times the probability that the player stays healthy plus \$0 times the probability that the player is injured.

However, players are generally more risk averse than teams. More is at stake for the player, as each negotiation represents all (or least the large majority) of a player's income, but only a fraction of the team's costs. The more risk averse players are, the more willing they are to sacrifice average yearly salary to guarantee

future income. This is a standard approach to a compensating wage differential. Players sacrifice compensation for the amenity of job security and the extent to which average yearly salary is decreasing in L depends on the degree to which players are risk averse. However, this assumption of relative risk aversion may not be realistic. MLB is a special case in which the bargaining power of the player can grow significantly over the course of his career. Kahn (1993) shows that contract length in MLB tends to grow as teams become concerned about a player entering free agency, and Maxcy (2004b) shows MLB owners have an incentive to mitigate risk by offering long-term contracts to players depending on market uncertainty as well as productive uncertainty. In this case, teams pay a premium for locking down players to longer contracts and players can expect increasing returns to L .

If the parties have similar expectations on the distributions of δ_1 and δ_2 , as will be the case if information is symmetric and complete, then they will choose F_j and W_i such that $A_{i,j}$ falls within the range of acceptable salaries. Here, they will complete negotiations privately. If one or both parties misinterpret the contract zone, causing them to take a stance that is too aggressive, then expectations of $A_{i,j}$ may fall outside of the range of acceptable salaries for both parties. If this is the case, an impasse results, and FOA is automatically triggered as an impasse resolution mechanism.

If this happens, player i and team j publicly declare target salaries, W'_i and F'_j , respectively. An arbitrator then selects one of these two salaries as a binding contract. The publicly-declared W'_i and F'_j may differ from privately held target salaries, as expectations of the arbitrator's valuation must be considered when framing an offer. Parties must incorporate the probability that they will lose the hearing into the terms they present to an arbitrator. Since a loss in FOA amounts to total concession to an opponent's terms, parties may be willing to reduce the surplus they extract from their term in order to increase their chances

of swaying the arbitrator. In the absence of arbitrator bias, W'_i and F'_j are arranged around the parties' expectations of the arbitrator's preferred award, which will be in the range of $A_{i,j}$ thereby resembling private negotiation. This arrangement is symmetric if both parties are risk neutral and information is symmetric. Parties must develop a strategy in an arbitration hearing that considers the benefit of the terms presented as well as the probability of the arbitrator's ruling in their favor. A party increases the benefit of an arbitration win with more aggressive terms, however, aggressive terms are less likely to be selected by an arbitrator. A more conciliatory stance increases the probability of winning the arbitrator's selection, but lowers the benefit of the win. Teams are better off as salary decreases while the opposite is true for players. Both parties increase their subjective probabilities of winning the hearing by moving closer to what they believe to be the arbitrator's preferred award. Given identical risk preferences and information, the parties' optimal positions will mirror one another around the expectation of the arbitrator's decision. If one party is more risk averse, its proposed salary will be closer to the expected arbitrator value.

In keeping with past literature, in which the average of proposed term is understood to significantly influence the arbitrator's criteria (Marburger and Burgess 2004, Marburger 2004), I borrow the framework for arbitrator behavior from Farber (1981) and apply it to an FOA framework. First, the arbitrator formulates a preferred salary figure, A_p^a . The arbitrator believes this figure to be equitable based on his a priori understanding of the negotiation impasse, which comes from available information about the parties and the market. Secondly, the arbitrator considers the salaries proposed by the player and the team, W'_i and F'_j respectively. Once parties state their cases, the arbitrator considers how close W'_i and F'_j are to one another, as well as which party presented stronger arguments. The relative strengths of the arguments is reflected in the variable $C_{i,j}$.

$C_{i,j}$ is negative if the arbitrator determines that the team's argument is stronger relative to the player's and positive if the player's argument is more convincing. This further borrows from literature suggesting that arbitrators do not consider the offers alone, but also the facts of the case presented (Bazerman and Farber 1985, Faurot and McAllister 1992). The magnitude of $C_{i,j}$ depends on the relative strengths of the arguments. Finally, the arbitrator uses all four elements, A_p^a , W'_i , F'_j , and $C_{i,j}$ to determine the salary figure, $A_{i,j}^a$, one which he bases his decision. In a CA framework, $A_{i,j}^a$ would be the resulting salary, but in FOA, the arbitrator uses this figure as a baseline to determine which party's terms to choose. The arbitrator rules in favor of the party that proposed terms closest to $A_{i,j}^a$. The arbitrator bases the usefulness (and thus the weighting) of A_p^a , W'_i , and F'_j on the proximity of W'_i , and F'_j to each other. If the parties' proposals are relatively close together, then the average of their offers holds more weight for the arbitrator's formulation of $A_{i,j}^a$. If the ask and the offer differ greatly, the arbitrator relies more heavily on his own determination of player value. Specifically, $A_{i,j}^a$ is a function of the arbitrator's preconceived opinion of the player's value, A_p^a and the average of the two parties' terms, such that:

$$A_{i,j}^a = (1 - \gamma)A_p^a + \gamma\left(\frac{W'_i + F'_j}{2} + C_{i,j}\right) \quad (2.4)$$

and

$$\gamma = h(W'_i - F'_j)$$

where $h(W'_i - F'_j)$ approaches one as $W'_i - F'_j$ approaches zero. Once the arbitrator has settled on equitable terms and received the bids, he must side with

one party or the other based on who is closer to $A_{i,j}^a$, such that:

$$A_{i,j} = \begin{cases} W'_i, & \text{if } |W'_i - A_{i,j}^a| < |F'_j - A_{i,j}^a| \\ F'_j, & \text{otherwise} \end{cases} \quad (2.5)$$

When negotiating average yearly salary, the parties must consider the arbitration environment, even if they do not employ an arbitrator. Bringing together the five equations outlined above shows average yearly salary to be a function of the parties' strategies in and out of the potential arbitration hearing. The combined considerations show AYS to be a function of expected player performance (U_i and V_i expected by the team and player respectively), length of contract (L), and the components of the environment created when an arbitrator is included ($A_{i,j}^a$, W'_i , and F'_j) such that:

$$A_{i,j} = q(U_i, V_i, L, A_{i,j}^a, W'_i, F'_j) \quad (2.6)$$

I identify the following null hypotheses for the above model. First, given the assumption that players are more risk averse regarding future bargaining power and potential injury, the coefficient on contract length should be negative. Risk averse players should be willing to reduce Average Yearly Salary for a stable income over more years. Even if teams are risk neutral, they discount future value, and thus are willing to offer less for additional contract years.

Secondly, when players and teams avoid arbitration, the average yearly salary more closely resembles the Ask than the Offer. This is, again, because players are more risk averse than teams are. Each negotiation represents a small portion of the team's total costs, as opposed to the player who is negotiating the entirety of his salary. Thus teams are more able to diversify away their risk. Risk

aversion manifests itself in two ways, risk of losing an arbitration hearing during current negotiations and risk of reduced bargaining power in future negotiations due to injury or decreased productivity. Greater risk aversion during a hearing motivates players to bring an Ask that is more conciliatory than the team's Offer to the arbitration hearing. Parties are trying to extract as much surplus as possible while still presenting terms that are closer than their opponent's to the value that they believe the arbitrator deems optimal. When a player reduces his Ask, he is paying a premium to make his bid more favorable in the eyes of the arbitrator and thus increase his probability of winning the hearing. The same can be said of a team's increasing its Offer. If players are more risk averse than teams, they are willing to pay a higher premium, by moving their terms closer to the arbitrator's expected estimation, in order to increase their likelihood of winning the hearing.

According to the Arbitrator Exchangeability Hypothesis, each arbitrator's assessment of player value should converge toward that player's marginal revenue product (MRP)⁷ If the arbitrator's preferred award regularly exceeds the player's true value, then he will be vetoed by teams in the future. Likewise, players will veto arbitrators that they expect to undervalue players. This leaves no opportunity for either party to take advantage of arbitrator bias. Therefore, as parties arrange their figures around their respective expectations of the arbitrator's preferred award, they are operating within a contract zone that includes a player's expected MRP. However, lingering uncertainty surrounding the arbitrator's predictability and asymmetric risk aversion suggests that, while the player's true expected MRP will fall between the team's Offer and the player's Ask, it is not necessarily a weighted average of the two figures.

⁷While different markets will have differing ability to generate revenue, MLB arbitrators are prohibited from considering ability to pay when making their decisions (Abrams 1999). MRP, in this context, can be viewed as the average value within the league, not the value added to specific team.

2.5 Data and Variables

2.5.1 *Data*

Data include exchanged figures between teams and arbitration eligible players. When teams and players exchange arbitration figures, they do so well in advance of the arbitration hearing.⁸ The player and the team are free to continue negotiating after this information is revealed, typically prompting a settlement before the hearing takes place. This provides insight into the parties' FOA strategy even if they avoid a hearing. Data frequently contain information on both FOA and private negotiation, allowing me to analyze whether FOA as an impasse resolution mechanism succeeds in its goal to yield outcomes that resemble private negotiation. This comparison shows whether arbitration promotes bargaining and whether the outcomes resemble expected outcomes in the absence of such an impasse mechanism. I have gathered data on all MLB arbitration-eligible players during the period 2007-2014 from Maury Brown at *The Biz of Baseball*. Game-play data is gathered from *Baseball Prospectus*.

2.5.2 *Variables*

The dependent variable is a player's average yearly base salary, AYS. This is determined by taking the total value of the payments and dividing it by the length of the contract. This normalizes the payout, though salary may be unevenly disseminated over the life of the contract. *Ask* and *Offer* refer to the official, exchanged figures announced by the player and team, respectively. Exchanged figures are one-year salary figures, and do not include bonuses or performance

⁸Despite this exchange happening well in advance of a potential hearing, not all negotiation reach this point. In some cases, parties settle on a contract early in the process, even before arbitration figures are announced.

incentives. In practice, each party's target salary (W_i and F_j) is unknown. However, players and teams commonly exchange official arbitration hearing figures that would be brought to a hearing. These are the only officially announced figures prior to a settlement. Once announced, these figures are locked-in, though private negotiations can continue after figures are exchanged. Should a hearing occur, the arbitrator will choose between these announced figures. This allows me to collect data on the player's Ask, W'_i , and team's Offer, F'_j , even if there is no hearing.

Length is a count variable indicating the duration of the contract in years. If negotiations go to an arbitration hearing, this equals one.

Most negotiations are settled privately, even after figures are exchanged. If private negotiations are not settled by the deadline, an arbitrator selects one of the two officially exchanged figures and writes it into a one-year, binding contract. Two dummy variables are included to indicate the outcome when a hearing occurs. PlayerWin equals one when a player wins his arbitration hearing, zero otherwise. TeamWin equals one when a team wins an arbitration hearing, zero otherwise. When arbitration is avoided, both of these variables equal zero. These outcomes capture information on an arbitrator's preferred award, $A_{i,j}^a$.

I include two lags of Wins Above Replacement Player (WARP), a metric developed by *Baseball Prospectus* (www.baseballprospectus.com, 2015) that captures overall player performance. I choose a lag of 2 because some players become arbitration eligible after only two years of service⁹. Thus, two lags is the most I can include without having to drop observations. I maximize observations because MLB is a relatively small organization, and the dataset, while comprehensive, is not large¹⁰. WARP captures teams' and players' expectations of future

⁹see footnote 6

¹⁰Including a third lag of WARP forced me to drop 13 of 316 observations. Additionally, I

performance (U_i and V_i). I choose WARP as the performance metric because it includes a wide array of contributions, including playing time, position, batting, base-running, and defense for batters, and role, innings pitched, and quality of performance for pitchers (*Baseball Prospectus 2015*). This allows me to capture offensive and defensive contributions with a single variable, thereby reducing the degrees of freedom I must sacrifice when controlling for performance. WARP is normalized such that one additional unit of WARP contributed by a player should increase overall team performance by one win. A performance of WARP=0 represents that level of contribution expected from a “replacement player”.

Replacement level identifies a level of talent high enough to contribute at the Major League level, but low enough that it is always available in a team’s minor league system. A particularly poor season by a player may result in a negative WARP. In these cases, the team would have been better off promoting a replacement level player. At the other end of the spectrum, a player’s WARP has exceeded 12.2 in a single season on just five occasions in MLB history; Sandy Koufax (1965 and 1966), Albert Pujols (2009), and Mickey Mantle (1956 and 1957). Table 1.2 (in chapter 1) is a cross correlation table that includes several popular baseball statistics used to measure player performance. These statistics are gathered at the team level and WARP is revealed to have a much stronger correlation with wins when compared to some of the more commonly used metrics.

Lastly, I include the dummy variable, Pitcher, which equals one if the player is a pitcher and zero otherwise. Pitchers tend to play less often than hitters, requiring extra days of rest between performances. A starting pitcher will typically get four days of rest following each outing. Relief pitchers are also given more frequent rest and are often called upon only in specific situations (i.e., a left-handed pitcher pitches to a left-handed hitter). However, even though pitchers appear in

will show that none of the lags of WARP are significant in the fixed effects model.

fewer games, whenever they are on the field they are necessarily involved in every defensive play. Shouldering a more intense workload with less frequency creates a different type of wear-and-tear on the body, which makes pitchers prone to different, perhaps more severe or career-altering, injuries than hitters. While WARP captures both offensive and defensive contributions, the nature of the pitcher position, coupled with less frequent outings on which to collect data, likely increases uncertainty surrounding pitchers making it more difficult to predict future performance for pitchers relative to hitters. This potentially provides less certainty in terms of player consistency and value.

2.6 Empirical Analysis

2.6.1 OLS and Fixed Effects Model

The regression model described above appears in Equation 2.7, and summary statistics are included in Table 2.2:

$$\begin{aligned}
 AYS_{t,i,j} = & \beta_0 + \beta_1(Ask_{t,i,j}) + \beta_2(Offer_{t,i,j}) + \beta_3(Length_{i,j}) + \beta_4(PlayerWin_{t,i,j}) + \\
 & \beta_5(TeamWin_{t,i,j}) + \beta_6(Pitcher_i) + \beta_7(WARP_{t-1,i}) + \beta_8(WARP_{t-2,i}) + \epsilon
 \end{aligned}
 \tag{2.7}$$

Table 2.1: Summary Statistics

Variable	Mean	Std. Dev.	Min	Max	N
Average Yearly Salary†	3.289	2.892	.5	30.7	1007
Ask†	4.14	2.94	.59	21.5	344
Offer†	3.10	2.25	.50	17.00	344
Length of Contract	1.178	0.754	0	8	1044
Player Wins Arbitration	0.014	0.119	0	1	1044
Team Wins Arbitration	0.023	0.15	0	1	1044

† in millions

As a baseline, I run Pooled Ordinary Least Squares with year and team fixed effects. Year fixed effects control for overall salary growth over the years in the dataset, and team fixed effects control for fixed differences in teams' ability to compensate players. As a robustness check, I also run a player fixed effects (FE) model in order to control for unobservable difference among players. These unobservable differences may include things like clubhouse presence, leadership, work ethic, sportsmanship, local fan relationship, or a unique bond with the team or coaching staff. These are elements that may factor into how proposed terms are negotiated, but are not quantifiable. These qualities are likely reflected in the Ask and Offer, but portions may remain in the error term, if they have been left to be sorted on the negotiation table. In the FE model, I drop team fixed effects, as players in the dataset do not tend to change teams. The vast majority of players in the dataset (greater than 88%) only ever become arbitration eligible with a single team, and the inclusion of team fixed effects creates a multicollinearity problem.

Table 2.2 shows results from the Ordinary Least Squares (OLS) and Fixed Effects (FE) regressions. The coefficients on the Ask and Offer variables are intuitive. If the parties privately negotiate a settlement midway between the exchanged figures, both coefficients equal 0.5. Results show that both coefficients are statistically indistinguishable from .5, and their sum is statistically indistinguishable from one. Theory suggests that if both the firm and worker are equally risk-averse, then the "offer" and "ask" will fall symmetrically around their expectation of the arbitrator's valuation, provided both parties have identical information about a the player's future value. Thus, it is unsurprising that the difference in the coefficients is not significant. Recall, however, that these are exchanged arbitration figures and not necessarily true target salaries. Teams and players seem to approach the hearing in a way that balances the weight of their respective

2.6.2 OLS and Fixed Effects Results

Table 2.2: Pooled OLS and Fixed Effects estimations, Factors Affecting Average Yearly Salary of Arbitration Eligible Players

VARIABLES	(1) OLS	(2) FE
Ask	0.510*** (0.0731)	0.466** (0.201)
Offer	0.453*** (0.0941)	0.490** (0.238)
Length of Contract†	1.22*** (.037)	1.49*** (.155)
Player Wins Arbitration†‡	.497415*** (190,490)	1.018* (596,480)
Team Wins Arbitration†‡	-.169070 (149,602)	-.109573 (483,286)
Pitcher‡	147,890* (75,352)	-
WARP _{t-1}	89,935*** (26,162)	119,732 (95,125)
WARP _{t-2}	29,486 (27,322)	-22,137 (102,692)
Constant†	-1.710*** (.194318)	-1.711*** (.472101)
Year Fixed Effects	yes	yes
Team Fixed Effects	yes	no
Observations	337	337
R-squared	0.970	0.967
Number of Players in Panel		286

† in millions, ‡ dummy variable

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

terms. This not does necessarily mean that terms are arranged around the player's expected MRP.

The first lag of WARP, is positive and significant in the OLS model. This is intuitive, as greater WARP indicates greater value. The second lag is

not significant in the OLS model, and neither coefficient is significant in the FE model. This lack of significance is not surprising, as the parties' understanding of past performance is likely factored into the terms presented. Once terms are exchanged and I control for player fixed effects, past WARP offers little additional information.

A positive coefficient on the dummy Pitcher indicates that teams pay a premium for pitchers, though the increase is significant at only the 10% level. The pitcher dummy is not included in the FE model, as there is no intertemporal variation in this variable. Only in extremely rare circumstances does a MLB pitcher give up pitching to continue a career as a hitter, or vice versa.

The signs of the coefficients on the two dummy variables indicating who won an arbitration hearing are also intuitive. The vast majority of negotiations do not make it to the hearing stage, in which case both dummy variables equal zero. It makes sense that the coefficient on a player win is positive, while the coefficient on a team win is negative. Results indicated that a player win adds, on average, approximately \$1 million to salary for the following year, relative to settling without arbitration. A team win in arbitration, on average, reduces salary by \$100,000 compared to a single season salary in which arbitration is avoided. The coefficient on team wins is not significant in either model ¹¹. This suggests that players have little to lose by entering arbitration. Teams, on average, approach arbitration with terms that are not meaningfully different from what could be expected from private negotiation, while players take a significant step forward if their terms are selected. This result may indicate that players who can command unusually high salaries are systematically different from players who expect unusually low salaries. This will be addressed further in section 2.6.3.

¹¹Though I will show later that it is significant in each quantile when I run a quantile regression.

The most intriguing result is the coefficient on Length of Contract, which is positive. This is consistent with the findings of Maxcy (2004b). If players are more risk averse than teams, one would expect that, as parties discount future value, teams would be willing to pay and players would be willing to accept less per year, if the total value of the contract grows by adding more years to the terms. This positive coefficient suggests an increasing marginal cost to the team of offering players extra years in their contracts, indicating that the teams are paying a premium to mitigate risk. Specifically, according to the FE model, a player can expect, on average, an additional \$1.49M added to his yearly salary, per additional year on the contract. This means a two-year contract, would yield an additional almost \$3M *per year*. This seems counterintuitive, unless one considers the possibility that the arbitration process is systematically biased against players. If players know that their salary will increase significantly once they break free of the arbitration system and enter free agency, then they will demand a higher salary for contracts that extend into this new system, when their bargaining power will improve.

One possible interpretation of the coefficients on Length of Contract is that a few players with particularly lengthy contracts, which bleed into the years of free agency eligibility, are skewing the results. This could be attributed to a metric, observable to teams but not quantified by the data, that indicates a player's durability and ability to play at a consistently high level over an extended period of time. To address this concern, I removed all contracts with a length of at least 4 years. A four-year contract guarantees that the team has purchased all years of arbitration eligibility and at least one year of free agency. Additionally, a Chow test indicates that regression results differ for players with contracts of four years or greater. This restriction leads me to remove 21 of the 344 observations leaving only players whose contracts cover a maximum of three years. If my hypothesis

is correct, and players who command lengthy contracts are systematically different from all other players, then removing them from the data should reduce the coefficient on Length of Contract.

Table 2.3: Contracts Removed from Sample as Robustness Check

Player	Ask	Offer	Length	Total Contract	AYS	AYS-Ask
Freddie Freeman	5.8	4.5	8	140	17.5	11.7
Chase Utley	6.3	4.5	7	85	12.14	5.84
Homer Bailey	12	8.7	6	110	18.3	6.3
Nick Markakis	5	2.9	6	62	10.33	5.33
Felix Hernandez	12	7.2	5	78	15.6	3.59
Jose Bautista	11	7.6	5	64	12.8	1.8
Justin Verlander	9.5	6.9	5	80	16	6.5
Craig Kimbrel	9	6.6	4	42	10.5	1.5
Martin Prado	7.1	6.7	4	40	10	2.9
Aaron Harang	5.5	4.3	4	37	9.25	3.75
Josh Johnson	5	3	4	39	9.75	4.75
Robinson Cano	4.59	3.2	4	30	7.5	2.9
Joe Mauer	4.5	3.3	4	33	8.25	3.75
Zack Greinke	4.40	3.4	4	38	9.5	5.1
Billy Butler	4.3	3.4	4	30	7.5	3.2
Ervin Santana	4.3	3.6	4	30	7.5	3.2
Brandon Phillips	4.2	2.7	4	27	6.75	2.55
Bill Hall	4.1	3	4	24	6	1.9
Johnny Cueto	3.9	3	4	27	6.75	2.85
Michael Brantley	3.8	2.7	4	25	6.25	2.45
Yadier Molina	2.8	1.9	4	16	4	1.2

All 21 removed players received an average yearly salary that was much larger than the player would have asked for in the single-season arbitration hearing. The largest discrepancy is Freddie Freeman’s contract, signed in 2014. Freeman officially announced an arbitration figure of \$5.75M, but he eventually settled privately for a 8-year, \$140M contract, approximately \$17.5M per year. The smallest discrepancy for a 4+ year contract is for Yadier Molina. He announced an arbitration figure of \$2.8M, but he and the team settled on a 4-year contract totaling \$16M. All players received deals with AYS larger than the figure they would have presented in a hearing, suggesting that these contracts were driving the positive

coefficient on the Length of Contract variable. As expected, when the 21 observations were removed, the coefficient dropped. However, the coefficient remained positive and significant at the 99% level, and only dropped from approximately \$1.25M to \$895K in the OLS model and to \$1.08M in the FE model. This suggests that teams pay a premium for extra years, even for players who do not command long-term contracts.

Table 2.4: Pooled OLS and Fixed Effects estimations with 4+ year Contracts Removed, Factors Affecting Average Yearly Salary of Arbitration Eligible Players

VARIABLES	(1) OLS	(2) FE
Ask	0.453*** (0.0625)	0.467** (0.184)
Offer	0.537*** (0.0803)	0.432* (0.217)
Length of Contract†	.895*** (.055)	1.08*** (.201)
Player Wins Arbitration‡	524,756*** (154,987)	399,325 (573,009)
Team Wins Arbitration‡	-255,714** (119,212)	-306,436 (438,757)
Pitcher‡	40,437 (61,888)	-
WARP _{t-1}	49,584** (21,610)	16,209 (90,271)
WARP _{t-2}	37,024* (22,363)	-56,609 (95,671)
Constant†	-1.185*** (.166114)	-1.246*** (.446887)
Year Fixed Effects	yes	yes
Team Fixed Effects	yes	no
Observations	316	316
R-squared	0.975	0.962
Number of Players in Panel		268

† in millions, ‡ dummy variable

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

2.6.3 *Quantile Regression Model*

Results shown in the previous section indicate that there is a difference between players who sign longer contracts and those who sign shorter ones. When I analyze the subset of data that includes only contracts of 3 years or less, I find coefficients on Length of Contract that are statistically different from those found when analyzing the full data set. Two factors are at play here. Firstly, teams have to weigh the benefits of entering into a longer commitment with a player against the increasing probability of injury to, or decreased performance from, that player over time. Secondly, teams need to consider how the negotiation environment and a player's bargaining power shift over time. Longer contracts are the result of a team's feeling that the risk of having to negotiate in the future under circumstances that are more favorable for the player is greater than the risk of the player's losing value due to injury.

In the previous section, I used a subset of contracts less than four years in length because a contract that covers more than three years purchases at least one year of free agency. However, it is unlikely that this analysis fully uncovers the differences between players with longer contracts and those with contracts for fewer years. Players may sign multi-year deals during their second or third years of arbitration eligibility such that even a two-year contract may cross the threshold of eligibility from arbitration to free agency. Additionally, players who command higher than expected salaries may be fundamentally different from players who settle for unusually small salaries.

In order to explore this difference in detail, I employ conditional quantile regression (CQR). Recent literature has shown the unconditional quantile regression (UCR), developed by Firpo and Lemieux (2009), provides results that are more appropriate for policy analysis (Maclean, Webber, and Marti 2014,

Rodriguez-Caro, Vallejo-Torres, Lopez-Valcarcel 2016). These studies show the benefit of UQR, yielding coefficients that indicate the effect of an independent variable on the τ -th quantile of a dependent variable, as opposed to CQR's results which speak specifically to the conditional distribution of the dependent variable. However, my question is motivated specifically by the conditional distribution. In the case of MLB, the negotiation environment is largely defined by the player characteristics outlined in the model, including whether the player is a pitcher, and lagged performance statistics. Unobservable characteristics are captured in the *Ask* and *Offer* exchanged. Conditional on these characteristics, teams and players operate within the confines of an expected distribution for *AYS*. At this point, the question becomes, if contract length is extended, does the conditional distribution of *AYS* shift in favor of either party? While CQR results do not allow me to speak to how a specific player might expect his salary to change given an additional year on his contract, it does allow me to address how parties might change a negotiation environment that is otherwise defined by player characteristics. I believe this is the more pertinent question for this research.

If a team fears that a player's bargaining position will improve in the future, then it is likely that the team would curtail this future expense by offering a contract that extends into free agency. Further, teams may view a player who is likely to receive an unusually high salary in the next year as a worker with the potential for rapid salary growth in future years. Salaries that exceed expectations likely stem from intangible factors such as leadership, clubhouse presence, ability to draw fans, or work ethic. In these cases, teams will be more aggressive in offering lengthy contracts in an attempt to induce the player's acceptance of a secure but discounted salary rather than face his demand for continued salary growth. Quantile regression allows me to compare how salaries for players in

higher conditional salary percentiles change as years are added to a contract in comparison to the change in low percentile salaries.

Past literature has shown that changes in bargaining structure alters the way that teams reward players for performance and that these changes depend on their level of pay. For example, Leeds and Kowalewski (2001) apply quantile regression to National Football League data to show changes in bargaining strength following the 1993 collective bargaining agreement. Similarly, I show how parties shift their strategies when they know a player is about to transition into a period of greater bargaining power. My focus is on how length of contract affects average yearly salary. I hypothesize that teams expect the players that can command higher salaries conditional on observable variables to gain the greatest increase in bargaining power over the course arbitration eligibility and into free agency. Therefore, I expect the coefficient on length of contract to be positive and increase in magnitude across quantiles, as teams spend more on a current contracts in order to protect themselves from a less favorable negotiation environment in the future. The functional form of the framework for this model is similar to the functional form employed in Section 2.6.1 with the addition of the independent variable, τ , which I use to represents the quantile.

$$A_{i,j} = q(U_i, V_i, L, A_{i,j}^a, W_i', F_j', \tau) \quad (2.8)$$

I continue to use AYS as the dependent variable. As independent variables, I continue to use two lags of WARP and a Pitcher dummy variable as controls for past performance, Length of Contract represents the total number of years covered under the resulting contract, and finally I use Ask, Offer, and dummy variables for arbitration hearing outcomes to capture arbitration strategy. I run the analysis first using quartiles. I start with such a coarse breakdown

because the dataset is small, and I want to establish a baseline before I explore with greater precision. Next, I run a second analysis using deciles. This allows for greater granularity and can be compared to the quartile results. Quartiles results are found in Table 2.5, and Tables 2.6 and 2.7 report decile results.

2.6.4 *Quantile Regression Results*

Table 2.5: Quartile Regression Results, Factors Affecting Average Yearly Salary of Arbitration Eligible Players.

VARIABLES	(1) q25	(2) q50	(3) q75
Ask	0.349*** (0.0598)	0.464*** (0.0225)	0.519*** (0.0377)
Offer	0.645*** (0.0601)	0.536*** (0.0287)	0.475*** (0.0456)
Length of Contract [†]	0.860*** (0.0840)	1.125*** (0.0534)	1.430*** (0.108)
Player Wins	152,069 (144,421)	375,516*** (100,079)	594,064*** (171,389)
Team Wins	-441,689*** (163,534)	-231,214*** (50,915)	-137,124*** (34,697)
Pitcher	24,655 (28,037)	23,800** (10,706)	1,082 (16,963)
WARP _{t-1}	24,870* (13,286)	9,513 (5,958)	8,937 (6,460)
WARP _{t-2}	17,894 (11,574)	7,294 (4,471)	-105.8 (4,993)
Constant [†]	-918,468*** (.104143)	-1.166*** (.057334)	-1.442*** (.107799)
Year Fixed Effects	Yes	Yes	Yes
Observations	337	337	337

[†] in millions

Standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Table 2.6: Decile Regression Results (A), Factors Affecting Average Yearly Salary of Arbitration Eligible Players

VARIABLES	(1) q10	(2) q20	(3) q30	(4) q40	(5) q50
Ask	0.372*** (0.0628)	0.308*** (0.0679)	0.423*** (0.0704)	0.462*** (0.0443)	0.464*** (0.0320)
Offer	0.587*** (0.0798)	0.693*** (0.0824)	0.572*** (0.0777)	0.538*** (0.0513)	0.536*** (0.0385)
Length of Contract [†]	0.470*** (0.113)	0.710*** (0.107)	0.946*** (0.0947)	1.001*** (0.0675)	1.125*** (0.0697)
Player Wins	132,536 (130,349)	197,540 (128,517)	311,919** (131,369)	297,974*** (92,147)	375,516*** (86,628)
Team Wins	-529,789*** (142,609)	-438,572*** (168,612)	-299,589 (195,894)	-252,046* (128,570)	-231,214** (95,421)
Pitcher	58,307 (78,732)	25,872 (35,329)	31,315 (21,640)	25,134 (19,363)	23,800 (15,483)
WARPt - 1	29,564 (35,174)	29,571* (15,819)	15,255 (12,140)	8,596 (7,544)	9,513 (7,778)
WARPt - 2	37,275* (19,006)	21,882 (14,042)	13,903 (9,088)	10,330 (7,694)	7,294 (7,138)
Constant [†]	-580882*** (.151508)	-784185*** (.120842)	-1.021*** (.098280)	-1.058*** (.073462)	-1.166*** (.075498)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	337	337	337	337	337

[†] in millions

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.7: Decile Regression Results (B), Factors Affecting Average Yearly Salary of Arbitration Eligible Players

VARIABLES	(1) q60	(2) q70	(3) q80	(4) q90
Ask	0.487*** (0.0332)	0.517*** (0.0292)	0.528*** (0.0315)	0.567*** (0.0507)
Offer	0.510*** (0.0383)	0.477*** (0.0337)	0.464*** (0.0369)	0.426*** (0.0553)
Length of Contract [†]	1.169*** (0.0993)	1.274*** (0.155)	1.470*** (0.134)	1.659*** (0.148)
Player Wins	385,106** (151,210)	557,718*** (171,862)	587,160** (234,588)	770,163** (313,858)
Team Wins	-150,472*** (20,485)	-147,019*** (22,146)	-123,923*** (32,813)	-110,451*** (34,128)
Pitcher	14,659 (9,830)	9,508 (11,982)	-635.8 (14,590)	9,066 (22,836)
WARPt - 1	11,852** (5,406)	10,297* (6,153)	10,488 (7,346)	16,984*** (5,358)
WARPt - 2	5,117 (4,622)	3,266 (3,834)	-810.8 (4,575)	-4,923 (5,669)
Constant [†]	-1.201*** (.103122)	-1.294*** (.154509)	-1.482*** (.129320)	-1.693*** (.147820)
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	337	337	337	337

[†] in millions

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

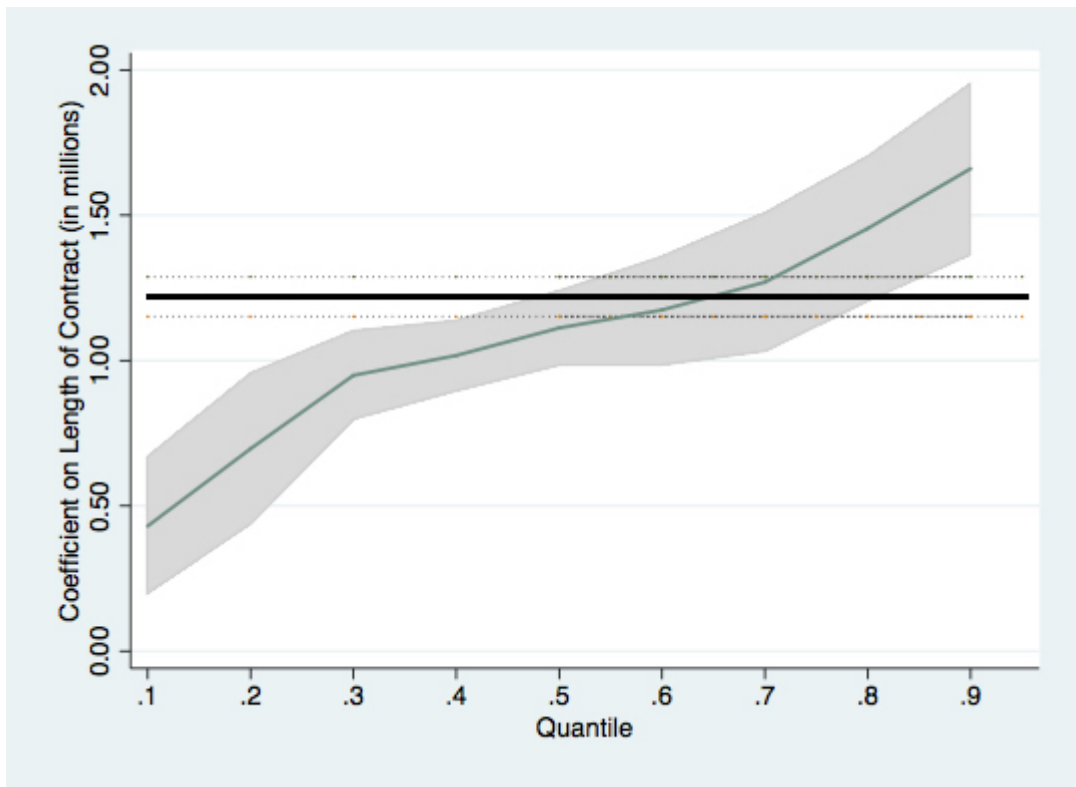
The above results show that contract length has a positive effect on each quantile. This means that as length of contract grows, salary rises for players earning unusually high salaries and unusually low salaries alike, though the increase in AYS is not of the same magnitude across quantiles. None of the quantiles indicate diminishing yearly cost for teams as would be expected if the arbitration process were assumed to yield outcomes that capture MRP. The coefficient on Length of Contract grows steadily through each quartile and through each decile. The 90th percentile estimate is statistically significantly higher than the median estimate. The median estimate is statistically significantly higher than the 20th percentile estimate, and so on. This indicates that arbitration-eligible players across the board stand to increase their AYS by negotiating additional years onto their contract, but the higher the conditional salary a player can command, the more that player can gain from negotiating an additional year. This makes length of contract an increasingly important factor during negotiations between teams and a players with higher than expected salaries.

A positive coefficient across all quantiles indicates that even at low conditional salaries, teams are paying a premium for longer contracts. Below is a graph of decile results relative to OLS results. The OLS coefficient on Length of Contract is the bold horizontal line, with the dotted lines indicating the standard error.

As shown, even the salaries of the bottom 10th percentile increase by almost \$500,000 per additional year. The 90th percentile increases by a full \$1.7M in AYS for an additional year. This likely means that teams, even when discounting future value, offer large increases in AYS in order to tie down players with multi-year contracts. This is counter-intuitive if one accepts the Arbitration Exchangeability hypothesis. This hypothesis assert that parties enter arbitration hearings with their respective offers arranged around expected MRP. However, assuming

that salaries will rise rapidly for players as they move through arbitration eligibility and enter free agency reconciles these results with intuition. In this case, buying extra years, even at a premium, may still be cost-effective for teams, if they believe the increased yearly salaries are lower than the salary increases one could expect from yearly renegotiations. If players are willing to accept a more conservative offer than what they might expect in subsequent arbitration hearings or the free agent market, then a rising cost for additional years may still represent teams' taking advantage of player risk aversion. Players who command top salaries conditional on the independent variables, are likely to see large bumps in salary in subsequent years. These players need larger incentives to forgo the open market, even when accounting for risk aversion.

Graph 1: OLS versus Quantile Results



Additionally, the quantile regression yields intriguing results for the dummy variables indicating the winners of arbitration hearings. Recall, that the dummy variable for team wins was insignificant in both the Ordinary Least Squares and Fixed Effects model. In the Quantile Regression model, however, the Team Wins dummy and Players Wins dummy are both significant up to at least the 90% significance level in every decile. The values of the coefficients on the Player Wins dummy are all positive, as expected, and increase in magnitude as deciles grow. This is the same in the quartile analysis, though the coefficient in the 25th quartile is not significant. Conversely, the coefficient on the Team Wins dummy is negative across quantiles but shrinks in magnitude, becoming less negative as deciles (or quartiles) grow. Winning an arbitration hearing appears increasingly important for players that command a salary that is particularly high conditional on observable variables. On the other hand, teams stand to gain the most by winning an arbitration hearing against a player that is expected to command an unusually low salary. If salary is unusually low, team's likely feel more comfortable taking an aggressive arbitration stance.

The coefficients on Ask and Offer are, unsurprisingly, significant at the 99% across all quantiles. The coefficient on Ask trends upward through deciles, while the coefficient on Offer trends downward as deciles increase. This makes intuitive sense, as it indicates that receiving an unusually high salary is the result of the player having strong bargaining power, and an unusually low salary stems from a team's strong bargaining power.

Lagged WARP is significant in the 25% quartile. It is not significant in any additional quartiles, nor is the second lag significant anywhere. In the decile analysis, a single lag yields results significant at the 99% level on the 90th deciles alone, while a second lag is significant at the 90% level in the 10th, 20th and 30th deciles. The significant coefficients are all positive, indicating that higher levels of

past performance helps a player to increase his pay. The coefficients are relatively small indicating that lagged WARP, even when significant, does not explain much variation in salary. This suggests that past performance has a slight impact only when players command unusually high or low salaries. Otherwise, as discussed for the OLS and FE models, performance is likely captured in the exchanged Ask and Offer.

Quantile regression results indicate that teams approach negotiation as if bargaining power will shift in player's favor in the future. This creates an incentive for teams to purchase additional contract years, even at a premium. Player that command unusually high salaries have the most to gain by selling additional years of service. Teams willing to pay this high premium for longer contracts likely believe that these are the players who stand to command the greatest increase in bargaining power in the future.

2.7 Discussion and Conclusion

The curiosity here is why the Player's Association would approve of a system that favors owners. Bias can come in a variety of forms. Arbitrators may consistently undervalue player performance, overestimate the likelihood of injury, or base their decision on past salary instead of true value. Similarly, as mentioned earlier, it is likely that arbitrators favor the team as the party that receives a lottery pay-off, meaning the player receives guaranteed salary as opposed to the team that receives expected production. Regardless, bias should be weeded out by the players union's ability to veto arbitrators. However, players who circumvent the arbitration process seem to highlight the notion that bargaining shifts in their favor for the contract years following arbitration eligibility. In order to justify

leaving this system unchallenged one must consider the goal of an arbitration system in the context of league-wide success.

The MLB arbitration system predates MLB free agency by two years, instituted in 1974 and 1976 respectively. Prior to the advent of arbitration, teams maintained the rights to the draftees' contracts in perpetuity. This suppressed player's wages significantly. However, it also allowed the league to maintain an incentive for teams to scout and develop young players without fear that those players, once trained and ready for MLB play, would quickly sell their services to other teams. While minor league development allows players to adjust to unique clubhouse cultures and coaching styles, a large portion of this training is also general job training. For teams to have an incentive to provide general training, they need assurance that they can retain the skills that they have paid to develop. If arbitration was intended to act as a compromise that facilitates salary growth for players, but also allows teams to keep the rights to player contracts, then it makes sense that some wage suppression would linger. Arbitration, in the case of MLB, was not meant to mimic free agency. After players successfully implemented a negotiation framework that included a pathway to free agency, the arbitration system remained intact. Players can still file for arbitration after three years of service, but now, in addition, they can file for free agency after six years of service. This allows arbitration to fill a role between the monopsony power of the reserve clause and the open market of free agency.¹² If, as theory suggests, arbitration yields unbiased, equitable settlements, then owners should be indifferent between granting arbitration eligibility and granting free agency. In fact, Marvin Miller, Executive Director of the Major League Baseball Players Association (MLBPA) from 1966 to 1982, suggested that the advent of the arbitration system has done more for

¹²This is analogous to Restricted Free Agency in other major sports leagues, in which the current team holds the "right of first refusal" wherein the current team is given the option to match a contract offer before the players signs with a new team.

players than free agency (Miller, 1991). However, even after this large shift in bargaining power, it appears that surplus is still being extracted from arbitration eligible players. The analysis in this paper is consistent with past literature suggesting that the increase in bargaining power when entering arbitration-eligibility is substantial, but a further shift occurs when entering free agency. The seminal work of Scully (1974) and well as Zimbalist (1992), Marburger and Burgess (2004), and Marburger (2004) all suggest that arbitration eligibility fills a gap, acting as a stepping stone during the bargaining power shift from newly drafted talent to free agency. Krautmann, von Allmen, and Berri (2009) indicate that the jump from arbitration ineligibility to eligibility causes the bargaining power to rise substantially for player. By their estimates, players are typically paid around 20% of their MRP prior to arbitration eligibility and over 85% of their MRP after earning eligibility.¹³ It is likely that small-market teams would be unable to afford the salaries doled out in arbitration and regularly forced to trade away these players, upsetting competitive balance.

The acceptance of the current system may be a bargaining tool used by MLBPA to leverage demands in other areas of negotiation. MLBPA regularly opposes wage suppression in the name of competitive balance. Teams, on the other hand, argue that if players could rely on an arbitration system that awards them their free market value, then small market teams would not be able to afford to keep their arbitration-eligible players, let alone compete in the free agent market.

According to theory, players could expect FOA to yield free market values so long as the union could veto arbitrators that were expected to rule below free market value. As a compromise between these two stances, arbitration may

¹³Percentages are based on the free market approach that assumes a competitive bidding process aligns free agent salaries with MRP (Krautmann, von Allmen, and Berri, 2009)

act as a middle-ground between complete monopsony power and free agency, but only if the players' union does not insist on vetoing arbitrators that consistently side with terms that are below players' free market value. An arbitration system that significantly increases player salaries but stops short of awarding a player his full market value seems likely to be a bargaining chip used to leverage a stronger position for players in the free agent market. It is true that arbitration greatly increases players' bargaining power, but it does not eliminate the exploitation of players entirely, as the Arbitrator Exchangeability suggests that it should. If arbitration eligible players still tend to make less than their MRP, then a team may sacrifice its current bargaining power in order to secure the player for future seasons before the player's bargaining power grows even more. This would reconcile the appearance that extended contracts are increasing in value with assumed player risk aversion.

The advent of the modern arbitration system broke down teams' monopsony power by limiting the reserve clause. This represented an enormous shift in bargaining power in favor of the players and undoubtedly increased players' salaries. However, the results of this paper's analysis indicate that teams approach contract negotiations as if salaries will increase significantly not only through the arbitration eligible years, but will continue to increase into the years of free agency. This provides an interesting insight into the Arbitration Exchangeability Hypothesis. Both teams and players consistently bring "Asks" and "Offers" to the table that are below the player's true value, as their expectation of an arbitrator's valuation is below the player's true value. This indicates that, even if arbitrators render decisions that are indistinguishable from those rendered by their peers, the arbitration system can still create an environment that suppresses wages relative to the free market. Wage suppression is revealed when we see parties shift their expectations to competitive market value forecasts when negotiating multi-years

contracts. The arbitration process maintains at least a portion of the muting effect that the reserve clause has on wages, and this bias against players is not fully shed until players reach full free agency. If the MLB arbitration environment is systematically biased against players, then the growth of average yearly salary as contract length extends may still represent a sum smaller than what a player could have expected to command in the open market.

Quantile regression provides more detailed insight into the bargaining power of players and teams. Teams' purchase of extra years comes at an added cost to teams across quantiles. In fact, in every decile, the coefficient on Length of Contract is larger than the coefficient on the dummy variable for a player's winning an arbitration hearing, indicating that for each contract, players see greater benefit from avoiding a hearing and negotiating a lengthy contract than they see from winning an arbitration hearing. This is likely due to a team's expectation that the player's salary will grow even faster if new contracts are negotiated year after year.

Additionally, quantile regression analysis shows a large shift in the importance of arbitration hearing outcomes for players and teams as players' salaries move from unusually low to unusually high. Players with unusually high salaries stand to gain the most by winning a hearing, while players with unusually low salaries see the largest decreases in salary when they lose a hearing. This indicates that players who win arbitration hearings have high salaries, conditional on observable variables, while the opposite is true when teams win arbitration hearings. This supports the hypothesis that teams offer longer contracts to players to whom they offer higher salaries. Further, this informs the risk for both parties associated with deferring to an arbitrator. The decision to privately negotiate or enter into an arbitration hearing involves careful consideration of not only player value, but of current and future bargaining power.

Despite appearing to be an unqualified win for players, the MLB arbitration system still allows teams to suppress salaries prior to free agency. Negotiating a multi-year contract and avoiding an arbitration hearing represents a compromise between the two parties. The approach of free agency shifts bargaining power in favor of players. This results in a positive relationship between average yearly salary and length of contract. Longer contracts that extend through arbitration eligibility allow players to capitalize on stronger future bargaining power. In exchange, teams are likely securing an expense that, while elevated early, still represents a future discount compared to renegotiation under less favorable conditions for the team.

CHAPTER 3

ARBITRATION AND UNCERTAINTY: AN EMPIRICAL ANALYSIS OF MAJOR LEAGUE BASEBALL'S NEGOTIATION LEARNING PROCESS

3.1 Introduction

When workers refuse to work (strike) or management refuses to continue production (lockout), a new sense of urgency can act as a catalyst to promote negotiation. In this way, both the threat and practice of strikes and lockouts are effective tools in overcoming a negotiation impasse. While potentially effective, whenever a work stoppage is actually carried out as a negotiation tactic, results are necessarily inefficient. The final agreement will be suboptimal, definitively dominated by an otherwise identical agreement negotiated prior to the work stoppage. Arbitration was designed as an impasse resolution mechanism that would avoid a work stoppage. If the parties cannot agree on the terms of a contract by a set deadline, they defer to a third-party arbitrator. The arbitrator presides over a hearing in which the parties present their preferred contract terms. In a conventional arbitration (CA) system, the arbitrator considers the respective terms

presented by each party as well as his own prior understanding of the dispute. The arbitrator then rules on what he believes to be equitable terms.

Final offer arbitration (FOA), formally introduced by Stevens (1966) is very similar to CA, except that the arbitrator is not free to formulate a compromise. Once terms have been presented, he must side with one party, whose terms are written into the contract. For the duration of this paper, I consider environments in which the arbitrator's decision is binding, for both CA and FOA.

This paper uses data from Major League Baseball (MLB) to empirically explore how the repeated use of the arbitration process over time influences subsequent negotiation. Additionally, I have arbitration data from the Wisconsin Employment Relations Commission (WERC). Both use an FOA system to resolve disputes between firm and worker. This can provide additional context as WERC data contribute not only information from another industry as a comparison to my results when using MLB data, but also represents a slightly different negotiation environment. Unlike MLB, in which teams and players arbitrate over individual salaries, WERC provides a context in which unions bargain over salary collectively. Further WERC employees arbitrate over non-wage amenities, such as health care, as well as wage. This is an important distinction as it very likely affects risk aversion. A union negotiator must concern himself with a wide array of constituents, while a sports agent need only worry about the risk aversion of a single client at a time. Further, professional baseball players possess an extremely specific labor skill-set, and yearly salaries are very high relative to the average American's yearly income. These different data sources allow me to consider the broader application of my results, and allows a glimpse into whether an industry as specific as professional baseball can provide insight into the FOA process in more general terms. As firms and workers continue to utilize the arbitration option, a learning process takes place that changes the uncertainty surrounding

the mechanism. This analysis sheds light on the behavior of negotiating parties, and helps to determine if arbitration succeeds in its goal of promoting private negotiation.

The paper proceeds as follows. Section 3.2 discusses the literature. Section 3.3 discusses discuss data. Section 3.4 provides a conceptual framework. Section 3.5 outlines an empirical model with results reported in section 3.6 reports results. Sections 3.7 and 3.8 discuss extensions to the work. Section 3.9 concludes the paper.

3.2 Related Literature

The primary focus of this paper will be the learning process of firms through repeated use of arbitration. The motivation for this analysis is to evaluate one of two criteria that Farber and Katz (1979) identify as necessary for the effective use of arbitration as a conflict resolution mechanism. Specifically, an arbitration system should promote private negotiation and should therefore be employed relatively infrequently. The other criterion is that, when an arbitration hearing occurs, the arbitrated settlement should resemble the outcome expected from private negotiation¹⁴. My model will build on past literature identifying necessary assumptions for arbitrator behavior that allow parties to reduce uncertainty surrounding an arbitration hearing through repeated use of the mechanism. Theory suggests that firms must consider any potential bias or personal incentives that arbitrator might have (Farber 1981, Farber and Bazerman 1986, Wittman 1986) as well as the direct costs of negotiation and arbitration (Bloom 1981).

Parties can, to an extent, facilitate this learning process by using the

¹⁴see Chapter 2 of this dissertation for a more detailed analysis of this criterion

veto process to control arbitrator behavior. I draw from literature on this theory of arbitrator selection in order to construct a model where firms and worker have an equal chance to curtail bias. Typically, both negotiating parties are given veto power over the choice of an arbitrator (Farber Bazerman 1986). The parties begin with a list of potential arbitrators and take turns crossing names off the list until the requisite number of names remain on the list as per the parties' collective bargaining agreement (CBA). This often means that a list will be narrowed to a single name, though some CBAs call for names to be crossed out until a panel of arbitrators remains. In these cases the hearing will be decided by a panel vote, rather than a single arbitrator. If an arbitrator develops a reputation for favoring the union's terms, then firms will veto him regularly. Similarly, unions will veto any arbitrator they perceive to have a bias in favor of firms. MLB represents a common arbitrator selection process wherein parties either agree privately to a panel of arbitrators, or jointly request a list of available arbitrators and then take turns crossing names of the that list (Faurot, McAllister 1992). In an effort to maintain consistent work, arbitrators balance their idea of objectively equitable terms with their own desire to garner favor from both parties.

This desire to be rehired is the basis for Arbitrator Exchangeability hypothesis (Ashenfelter 1987). If parties are given symmetric information, they will learn enough about the arbitrator's preferences to arrange their offers such that they are equally attractive to the arbitrator. This pushes all arbitrators toward an unbiased homogeneity, in which case any individual arbitrator is "exchangeable" for any other. This convergence is noteworthy as it will influence the uncertainty that parties face and the costs they associate with exercising an arbitration option.

My aim is to explore how parties are able to remove uncertainty within the negotiation environment by learning each time the arbitration mechanism is

triggered. Understanding the role of uncertainty in FOA is essential to understanding the learning processes of negotiating parties. Arbitration is an ineffective impasse resolution mechanism if hearing results are a foregone conclusion (making the deferral to an arbitrator a dominant strategy in all cases) or wildly unpredictable (thereby failing to resemble privately negotiated settlements).

The arbitration system avoids work stoppages, but uncertainty surrounding the arbitrator's decision creates a risk for both parties, and the potential cost should incentivize private negotiations. These costs are both direct and indirect. The direct cost comes in the form of payment to the arbitrator. This cost is shared by both parties and can be avoided entirely if terms are privately negotiated. Indirect costs to party_{*i*} are derived from the probability that the arbitrator's decision yields binding terms that favor party_{*j*}(*j*≠*i*) to such an extent that party_{*i*} is worse off than if it had acquiesced to the opposing party's terms during private negotiations. For example, Burgess and Marburger (1993) build on the theoretical literature with an empirical study using data from Major League Baseball (MLB). They find that FOA cases won by management typically yield salaries that are 9% lower than for comparable players who negotiated their salaries privately. The difference in outcome is statistically significant when controlling for differences between pitchers and position players, denoted hitters. Similarly, FOA cases that are won by players are 14% higher for pitchers and 4% higher for hitters than for comparable players who negotiated their salaries privately, though the increase for hitters is not statistically significant. This contrasts with previous empirical findings that FOA leads to more compromise during private negotiation (Bloom, Cavanagh 1984, Ashenfelter, Bloom 1984) including analysis specifically using MLB data (Coleman, Jennings, McLaughlin 1993). Empirical analysis also shows that CA, as opposed to FOA, in the National Hockey League leads to high quality settlements (Lambrinos, Ashman 2007). Still, this risk of an offer falling

outside of the private negotiation range may be helpful in promoting private negotiation. Marburger (2004) and Marburger and Burgess (2004) indicates that arbitrator preferences may be influencing the bargaining strategies of teams and players, and that parties can gain an advantage by learning from past decisions before formulating terms.

Babcock and Taylor (1996) make use of an experiment conducted by Del’Omo (1987, 1989) to create a measure of unpredictability. Their results show that an increase in uncertainty reduces the use of arbitration. This contradicts highlights the complexity of the role of uncertainty in arbitration: to promote private negotiation, one should employ unpredictable arbitrators, but unpredictable arbitrators may produce low-quality settlements. Here, “low-quality” means that an arbitrator may select an offer outside the expected range of privately-negotiated settlements.

Deck and Farmer (2006) explore the role of uncertainty further by developing a theoretical model, in which arbitration shrinks the contract zone. They further design an experiment to gather data and test the model yielding results that suggest arbitration discourages private settlement.

While weeding out arbitrator bias with a balanced veto process leads to higher quality settlements, it also affords parties much clearer insight into the arbitrator’s decision processes and therefore reduces uncertainty about expected arbitration outcomes. Specifically, I will explore whether this reduced uncertainty translates to fewer instances in which the mechanism is triggered, or if disagreements fall in magnitude when hearings do take place. The parties balance their desire to maximize returns with an effort to concede just enough to win the hearing; a decision that is directly influenced by the level of uncertainty and integral to my analysis. The parties’ behavior will change as they learn about the arbitration process through repeated use. Because arbitrators cannot secure work unless their

decision making process appears unbiased, lingering uncertainty must come from random shocks. This, however, assumes that collecting information is costless. In practice, it may be very costly to gather information to control for differences in each dispute and identify subtle underlying biases. However, potential gains from uncovering an arbitrator's bias incentivizes the parties to dig for information until the gains from acquiring additional information are outweighed by the costs at the margin. This process should drive arbitrators toward exchangeability, meaning arbitrators are indistinguishable from one another, but to what extent is this digging effective?

My analysis in the following section will add to this literature by contrasting two industries which appear to be learning how to avoid the arbitration process over time. Both MLB and WERC use final offer arbitration. FOA has been shown, specifically in MLB, to create an environment in which the proposed terms from teams and players play a vital role in defining the negotiation environment (Faurot McAllister 1992, Miller 2000). By developing a metric for the magnitude of impasse reach using the player's *Ask* and team's *Offer*, I explore how impasse changes over time, given repeated use of arbitration. This provides an stepping stone to analyzing WERC negotiation where terms involve more complicated compensation packages.

3.3 Data

I gather data on MLB arbitration hearings from the system's inception in 1974 through 2014, and I organize a panel by team¹⁵. Several time-invariant effects impact teams differently. Local substitutes for baseball entertainment, how conducive the local climate is to enjoying outdoor entertainment, and ease of access

¹⁵There are currently 30 MLB teams, but I have 31 unique team identifiers, as I created a new identifier for the Washington Nationals when the franchise moved from Montreal in 2005.

to the stadium all change in a new city. These differences affect local fan support and team revenue. This, in turn, affects the bargaining power of the team. The panel is unbalanced because of the Montreal Expos move to Washington DC as well as the fact that MLB went through 3 expansions over the data set's timespan. The data set consists of the number of arbitration cases heard for each team, the terms presented, and the winner of the hearing.

From the terms presented, I develop a metric, which I call the Disagreement Factor. This metric measures the magnitude of the negotiation impasse by dividing the distance between the terms presented by their average. This allows me to view the difference in the terms presented by teams and players relative to the approximate salary range in which negotiation is taking place. Specifically, I compute the metric by dividing the difference between the *Ask* and *Offer* by the average of the *Ask* and *Offer*. If the total difference between the *Ask* and the *Offer* is the "disagreement," the Disagreement Factor takes the extra step of expressing the disagreement relative to the range of proposed salaries. The total dollar value of the disagreement is not enough to determine the severity of the impasse. The disagreement needs to be viewed in the context of the general value of the worker, which is gleaned from the salary range in which parties are negotiating.

For example, consider a scenario in which a firm brings an *Offer* of \$10 and a worker an *Ask* of \$30, the disagreement factor would equal 1. That is, $(30-10)/20$. Now consider a second scenario in which the *Offer* and *Ask* are \$90 and \$110 respectively. In both cases, the total difference between terms is \$20, but in the second scenario, the disagreement factor has dropped to .2, indicating a less severe disagreement. While the disagreement is \$20 in both of these cases, the magnitude of the impasse falls as negotiations take place within higher and higher ranges.

I use the term Total Team Disagreement factor to refer to the sum of disagreement factors for all hearings heard by a single team in a given year. I also take the average disagreement factor for each team in each year. Summary Statistics appear in Table 3.1:

Table 3.1: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Total Team Disagreement Factor	1138	0.149	0.309	0	3.230
Average Team Disagreement Factor	1138	0.099	0.181	0	1.009
Team Cases Heard	1138	0.446	0.882	0	9
MLB Cases Heard	1138	12.014	8.605	0	35
MLB Cases Won by Players	1138	5.054	4.221	0	15

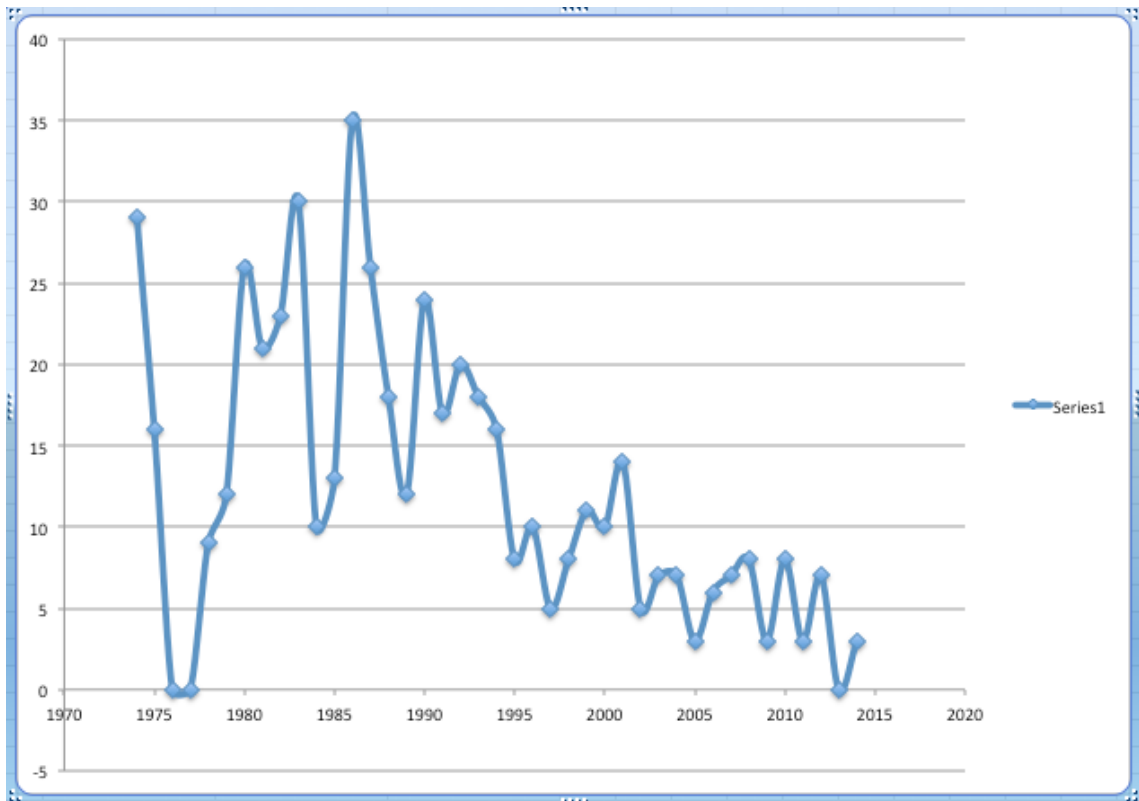
A mean Average Disagreement Factor of .099 means that the parties are, on average, arguing over approximately 10% of the average of their terms. To put this into context, consider that the median income in the USA in 2015 was about \$55,000. If a worker and firm were arguing over a salary that was expected to land around the median income, a disagreement factor of .099 would amount to the worker arguing for a salary of \$57,723 and the firm offering \$52,277. In contrast, the maximum average disagreement factor to appear in my dataset is 1.009. This would arise from a similar situation concerning the midpoint of the salary terms, but with the worker arguing for a salary of \$82,254 and the firm offering \$27,252. The Total Disagreement Factor relates less to individual cases, as it is the sum of disagreement factors. This is a factor of not only the extent to which parties disagree, but also the number of hearings.

Under an FOA system, arbitrators cannot “split the difference” between proposals. This raises the question of whether past rulings influence current actions by the parties. One way that the rehire incentive can influence an arbitrator is if a ruling in favor of a firm increases the probability that the arbitrator will rule in favor of a union at the next hearing, and vice versa. Arbitrators may worry about consecutive rulings in one direction, as they begin to fear a veto from the

side that has consistently lost, even if these losses were heard by another arbitrator. If the arbitrator is prone to placating a side that has lost in the past, even if its offer is inferior, then bias has been introduced into the subsequent settlements. In this context, “inferior terms” are farther from the arbitrator’s unbiased determination of an equitable settlement than the opposing party’s terms. When an arbitrator shifts his preferred settlement away from a party that has won in the past in an effort to appease the party that has lost in the past, he is basing the dispute outcome on information that is irrelevant to the dispute at hand. If this bias, real or perceived, is easily identified, then it can be exploited by the negotiating parties. As firms and unions become more familiar with potential arbitrators, uncertainty begins to wane. If the parties feel more sure of how the arbitrator will rule, they will feel less inclined to incur the cost of a hearing. The arbitration option will be used less, but will still influence the negotiation process.

The Arbitrator Exchangeability Hypothesis (Ashenfelter 1987) suggests that bias stemming from the rehire incentive will be weeded out by the veto process. However, even if arbitrators are assumed to be exchangeable, the negotiating parties may still have room to learn about arbitrators’ decision making processes. If so, they will glean this information by researching past decisions, as well as their own repeated use of arbitration hearings. If this is the case, the level of uncertainty changes over time. Figure 1 shows the total number of arbitration hearings by year. Note that the 1976 MLB collective bargaining agreement put a hold on arbitration for 1976 and 1977 while a free agency system was put into place. 1978 marked the reintroduction of arbitration and the three-tiered system still used today (Pappas, 2003). There is a clear downward trend in the number of cases heard. This indicates that the parties are privately settling with more frequency and finding it easier to avoid arbitration. This graph does not capture the magnitude of the impasses that continue to warrant the use of an arbitrator.

Figure 3.1: Total MLB Hearings 1974-2014



3.4 Conceptual Framework

During FOA, a firm and worker bring their respective terms before an arbitrator. The arbitrator cannot propose a compromise between the terms, and, instead, must rule in favor of one party or the other. The winning party's terms are then written into a binding contract. Since I am using MLB data, I will refer to the firm as the Team and the worker as the Player. The terms of the Team and Player will be denoted the *Offer* and the *Ask*, respectively. Since players move through the arbitration process for only a fraction of their careers, I approach this model from the perspective of the team, and its effort to negotiate with various players over time.

Focusing on the team's perspective still allows me to explore whether recent and repeated use of an arbitrator incentivizes private negotiation by pushing the both parties toward compromise. Specifically, if the team is the only party moving toward compromise, then it will win a disproportionate percentage of arbitration hearings. If teams are learning from the process and making an effort to reduce conflict, then players must respond in turn or risk entering negotiation at a disadvantage. Bearing this in mind, team behavior and negotiation results can also reveal player behavior. If players (or more likely, their agents) are learning at a similar rate to teams, then disagreements should dwindle, with no real advantage appearing in the probability that an arbitration hearing is actually won.

I use the disagreement factor in the model because a fixed difference in proposals is less severe the larger the size of the proposals. If players and teams are negotiating around a relatively high expected salary, then a fixed dispute represents a minor disagreement. If that same dollar value difference were to arise when negotiating a much lower expected salary, the dispute represents a much more significant conflict. For the lower salary range, the difference between the *Ask* and *Offer* will represent a much larger percentage change.

This metric builds on theory introduced by Farber (1979), wherein an arbitrator bases his decision on the *Ask*, *Offer*, and his own concept of an equitable award. The weight given to these variables is determined by the extent to which the parties agree, with greater weight given to the parties' terms the closer they are to one another. The arbitrator gives his own notion of an equitable settlement less weight if the parties agree to a large extent. The extent to which parties agree affects the arbitration environment.

I develop a model wherein the extent to which teams disagree with their players during arbitration hearings is a function of what they have learned from their own past experiences in arbitration hearings, as well as what teams and

players have learned when arbitration hearings are held throughout the league. Thus, the disagreement factor for team j in year t ($DF_{j,t}$) is a function of the team's disagreement factor from the previous year, the number of cases in which the team participated in the previous year, $cases_{j,t-1}$, as well as the cumulative number of cases heard across all MLB teams as of the previous year, $MLBCases_{t-1}$. The lagged disagreement factor is important, as it shows the team's approach in the most recent negotiation environment. This represents the most current and meaningful information from which the team can learn before the next season's round of contracts. The previous year is used for cumulative cases because all arbitration cases are heard within a relatively short window during each off-season. The previous year's cases are the most recent cases that all teams and players will have had time to year and consider, before entering the time frame to engage in the current year's cases.

$$DF_{j,t} = \beta_0 + \beta_1 DF_{j,t-1} + \beta_2 cases_{j,t-1} + \beta_3 MLBCases_{t-1} + \epsilon \quad (3.1)$$

Using $DF_{j,t-1}$ allows me to place extra weight on the information that the team has gathered most recently from its own experiences. I hypothesize that teams learn from each use of an arbitrator. All else equal, on average, the disagreement this year should only be a fraction of the disagreement last year because players and teams are both learning to approach hearings with more accurate expectations of the arbitrator's preferences. Therefore, I expect β_1 to be less than 1. The variable $cases_{j,t-1}$ accounts for learning from hearings regardless of the disagreement factor. If teams learn from each case, then the coefficient on $cases_{j,t-1}$ should be negative. Each case allows firms and teams to gather information that allows them to better prepare for a hearing. This reduces disagreements going forward. $MLBCases_{t-1}$ captures the way teams learn from all other teams in the league. Lastly, I include year fixed effects.

3.5 Empirical Model

My empirical model aims to estimate the magnitude of disagreements for a team in a given year, and whether that magnitude is reduced as parties learn from past experience. To begin, I use $TotalDF_{j,t}$ as my dependent variable. This is the sum of the disagreement factors calculated for each hearing in which team j participated during year t . $TotalDF_{j,t}$ does not distinguish between teams that have many arbitration hearings with relatively small disagreement factors and teams that have few hearings with large disagreement factors. To address this, I will repeat the analysis using an average rather than a sum. My initial empirical model is as follows:

$$TotalDF_{j,t} = \beta_0 + \beta_1 TotalDF_{j,t-1} + \beta_2 Cases_{j,t-1} + \beta_3 MLBCases_{t-1} + \gamma_t + \epsilon \quad (3.2)$$

First, I run a fixed effects estimator. Arbitration-eligible contracts are typically negotiated during the off-season and are all announced within a fairly short window. This prevents me from designing the panel with observation frequency greater than yearly. This estimator omits the lagged dependent variable. If unobserved panel-level effects are correlated with the lagged dependent variable, as is likely the case, then the inclusion of the lagged dependent variable introduces dynamic panel bias (Nickell 1981). Unobservable shocks linger in the lagged term causing correlation between the regressor and the error term which makes the estimator inconsistent (Roodman 2006).

However, if $TotalDF_{j,t-1}$ is a significant predictor, then dropping it in the fixed effects estimator likely replaces dynamic panel bias with omitted variable bias. In order to address this bias, I use an Arellano Bond estimator with robust standard errors. This GMM estimator corrects for dynamic panel bias by

using second (and further) lags of the dependent variable as well as lags of the independent variables as instruments for the lagged dependent variable in the first differenced equation. The results from both estimations can be found in Table 3.2 in the next section. Both models include year fixed effects to capture year specific changes such as renegotiated collective bargaining agreements, and market factors that influence the popularity of baseball entertainment and thus team revenue (steroids era, players' strikes, viewer-friendly World Series match-ups, etc.)

3.6 Results

Table 3.2: Fixed Effects and Arellano Bond Estimation of Total Disagreement Factor

VARIABLES	(1) FE	(2) AB
TotalDF _{<i>j,t-1</i>}		0.271** (0.111)
Cases _{<i>j,t-1</i>}	0.0725*** (0.0196)	-0.0392 (0.0528)
MLBCases _{<i>t-1</i>}	-0.00175 (0.00399)	-0.000281*** (.0000717)
Constant	0.903 (2.001)	0.131*** (0.0337)
Year Fixed Effects	Yes	Yes
Observations	1,107	1,076
R-squared	0.199	
Number of Teams	31	31

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Differences in the results from the fixed effects and Arellano Bond estimators provide evidence that Arellano Bond is the more appropriate estimator. When considering the high significance of $\text{TotalDF}_{j,t-1}$, in the Arellano-bond results, I can conclude that the fixed effects model is likely suffering from omitted variable bias by not including it in the estimation. This omitted variable bias is likely driving the counterintuitive positive sign of the coefficient on $\text{Cases}_{j,t-1}$. A positive coefficient unrealistically suggests that teams become less agreeable the more times they go through an arbitration hearing. If this were the case, team would not only be refusing to learn from past experience, but actively engaging in impasses that grow in intensity the more they participate in the process. Additionally, this result is not being driven by a particularly stubborn team culture, as team fixed effects control for such anomalies. Since the lagged Total Disagreement Factor is constructed by summing the disagreement factors of each case, I expect it to be positively correlated with the lagged total number of cases. Additionally, lagged Total Disagreement Factor has a positive, though fractional, coefficient in the model. If lagged Total Disagreement Factor is removed as a regressor to avoid dynamic panel bias, as the fixed effect model does, then I can expect the coefficient on lagged cases to be biased upward. This explains the positive coefficient on lagged cases in the fixed effects model, and the negative coefficient in the Arellano Bond model. This counterintuitive result from the fixed effects estimator supports my belief it is biased, and thus the Arellano-Bond estimator should be preferred.

As expected, the Arellano Bond estimation results return a coefficient on $\text{TotalDF}_{j,t-1}$ that is less than one. The negative coefficient on the number of cases heard by the team in the previous year is also expected, though not significant. This indicates that the disagreement factor declines with repeated use of the arbitration mechanism. Similarly, as teams participate in more cases,

they learn to moderate their proposals when entering hearings, which leads to cases with impasses that are less and less severe. $TotalDF_{j,t-1}$ is significant at the 95% level. $MLBCases_{t-1}$ is also significant (99% level) and negative. These results indicate that repeated use of arbitration reduces the magnitude of total disagreement. Over time, teams learn how to approach arbitration in such a way that the severity of total impasses are reduced. This likely reduces uncertainty and indicates that parties maintain more control over the resulting outcomes.

I expect $TotalDF_{j,t}$ to fall as teams learn what to expect from an arbitration hearing. However, using the Total Disagreement Factor does not identify whether teams are reducing the disagreement factor for each hearing, or simply engaging in fewer hearings. To explore this question, I also estimate a similar model, replacing Total Disagreement Factor with Average Disagreement Factor, $AVGDF_{j,t-1}$, as the dependent variable.

I calculate $AVGDF_{j,t}$ by taking the sum of team j 's disagreement factors from each hearing in year t , and dividing it by the number of hearings in which the team participated. Here, I interpret coefficients on the explanatory variables as the effects on an expected individual arbitration hearing, rather than a team's total disagreements across all hearings. Considering both specifications allows me to consider whether teams and players are learning how to reduce the disagreement factor by avoiding hearings, or by entering hearings with smaller relative differences in proposed salaries. As in the model in which $TotalDF_{j,t}$ is the depending variable, when using an Arellano Bond estimation, the coefficient on $AVGDF_{j,t-1}$ is positive and less than one, but here it is no longer significant. $Cases_{j,t-1}$ maintains a negative coefficient, but also remains insignificant. This is not surprising, as total cases understandably have a stronger correlation with total disagreement than with average disagreement. If teams are learning how to reduce disagreement and thus reduce the number of cases, the the learning process

will have a stronger impact on the total disagreement than the average disagreement. The negative coefficient on cumulative MLB cases supports the notion that teams learn from cases heard across MLB though the coefficient is practically small. Taken into consideration with Table 3.2, the results in Table 3.3 suggest learning about the arbitration process manifests itself in fewer hearings, though the hearings are not necessarily over smaller disputes when they do occur.

Table 3.3: Fixed Effects and Arellano Bond Estimation of Average Disagreement Factor

VARIABLES	(1) FE	(2) AB
AVGDF _{<i>j,t-1</i>}		0.0768 (0.0564)
Cases _{<i>j,t-1</i>}	0.0195*** (0.00642)	-0.0162 (0.0146)
MLBCases _{<i>t-1</i>}	-0.00327 (0.00280)	-0.000153*** (4.98e-05)
Constant	1.662 (1.408)	0.0768*** (0.0240)
Year Fixed Effects	Yes	Yes
Observations	1,107	1,076
R-squared	0.153	
Number of Teams	31	31

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

3.7 The Player's Perspective

One concern about my results is that they do not indicate whether players are learning as well. Since players are eligible for arbitration only during a very brief period in their careers, and many players negotiate privately to avoid arbitration hearings, an analogous panel for players would be extremely unbalanced and impractical. I provide some insight into this question, though, by considering a conceptual framework in which the choice of a winner in a hearing is a function

of the disagreement factor and the aggressiveness of the team's and the player's proposed salary such that:

$$Outcome_{t,i,j} = f(DF_{t,i,j}, Ask_{t,i}, Offer_{t,j}) \quad (3.3)$$

Outcome is the dependent variable indicating the side with which the arbitrator sided during the hearing held at time t , between player i and team j . *Ask* and *Offer* represent the respective proposed salaries from player i and team j at time t . If the aggressiveness of teams and players contribute equally to the disagreement factor, then the outcome from an unbiased arbitrator should be a coin flip. However, if one party is taking a particularly aggressive stance with its salary proposal, thereby driving the disagreement factor, then the probability that the arbitrator views its proposal favorably declines.

To test this hypothesis I use a binary dependent variable, *PlayerWins_{t,i,j}* which takes a value of one to indicate that the player won the arbitration hearing and zero otherwise. I use the disagreement factor as outlined previously, as well as the player's *Ask*. I need not include the team's *Offer*. Since I include the disagreement factor and *Ask*, and since the disagreement factor is derived from the *Ask* and *Offer*, the additional inclusion of *Offer* as an independent variable would be redundant. Additionally, I include a time trend, so the empirical model appears as follows:

$$PlayerWins_t = \beta_0 + \beta_1 DisagreementFactor + \beta_2 Ask + \beta_3 Trend + \epsilon \quad (3.4)$$

Ask is the player's proposed terms, which acts as a proxy for the salary range and aggressiveness of a player's negotiation stance, and *Trend* is a count variable that captures any systematic change in the probability of winning over time. My data include all 502 MLB arbitration hearings from 1974-2014. All

salary proposals have been adjusted for inflation. Dummy variables are used to identify teams in order to control for time invariant fixed effects. The results of a probit model are in Table 3.4.

Table 3.4: Probit Estimation of Arbitration Hearing Outcomes

VARIABLES	(1) Probit Marginal Effects
Disagreement Factor	-.0801 (0.135)
Player Ask [†]	5.68 (0.0000128)
Trend	-.00312 (0.0033724)
Team Fixed Effects	Yes
Constant	Yes
Observations	502

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

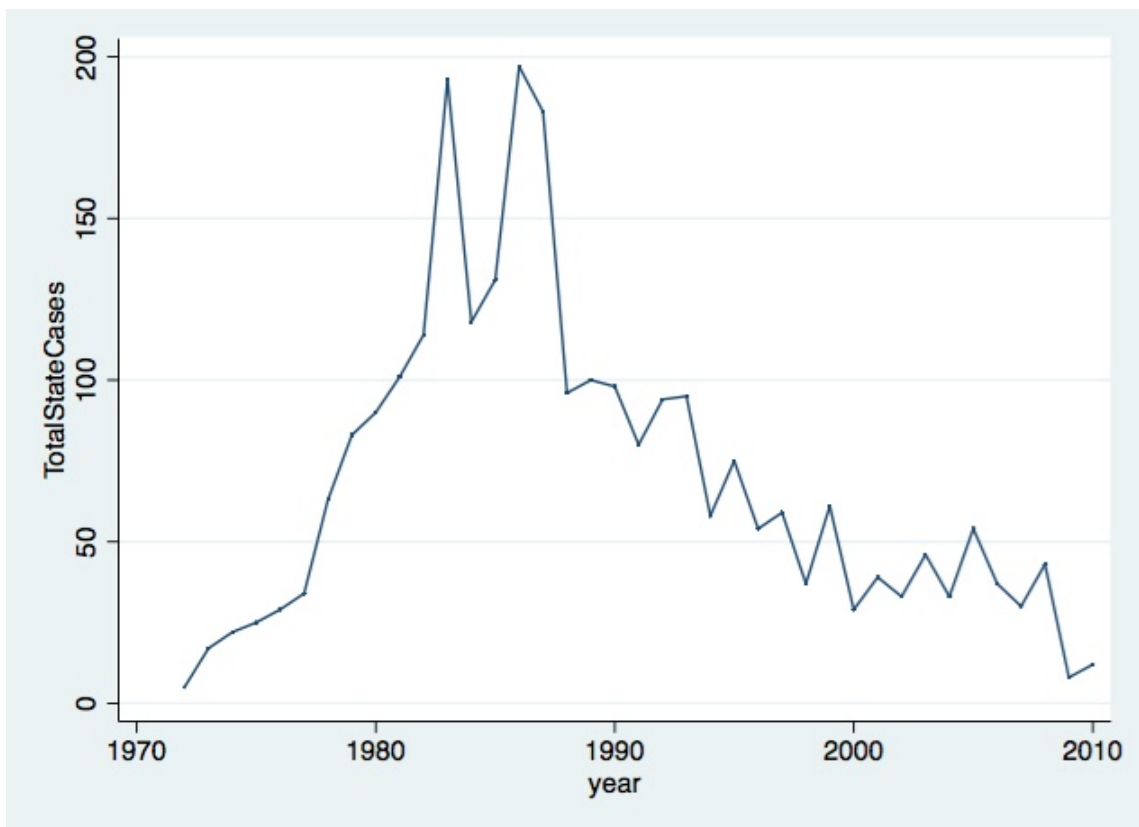
† in millions

According to these results, the disagreement factor is not a significant predictor of who will win a hearing. Thus, if the disagreement factor is shrinking over time, but neither teams nor players are gaining an advantage from this convergence, it is likely that both sides are learning at a similar rate. In contrast, if the disagreement factor were to have a significant and negative impact, then a shrinking disagreement factor gives teams an advantage. The reduction of impasse would be a tool used by teams with a better understanding of arbitrator preferences. An advantage to players would be indicated by a positive and significant coefficient. Even when the system fails to promote private negotiation, parties are exercising more control over outcomes by presenting terms that increasingly similar to one another.

3.8 Comparing Major League Baseball and the Public Sector

Data from the Wisconsin Employment Relations Commission indicates a similar trend. Like MLB, Wisconsin instituted an FOA system in the early 1970s. Unlike MLB, Wisconsin unions negotiate terms with their respective cities over many non-wage amenities in addition to the wage. Further, the hearings impact the entire union, as opposed to a single worker. Still, shown above, the number of FOA cases heard per year (1972-2010), follows a trend very similar to MLB, despite coming from a very different industry.

Figure 3.2: Total WERC Hearings 1972-2010



The data show what appears to be a clear learning process. In the early 1970s, FOA was used as an impasse procedure for Fire and Police Departments. By 1978 it had been adopted as the impasse procedure for the entire public sector. The number of hearings grows as the new procedure becomes the norm, but hearing volume clearly peaks in the mid to late 1980s and falls in the following decades.

I create another framework to determine if unions or governments are learning how to gain an advantage in arbitration hearings. Unlike, MLB, WERC bargains over much more than a single salary figure. Trades-offs between monetary compensation and non-wage amenities as well as proposed pay structures for multiple positions make the terms presented too complex for me to use the disagreement factor that I used in the MLB analysis. However, the number of hearings seems to be shrinking over time. This suggests that the parties are learning how to avoid arbitration hearings, and are thus reducing conflict, by negotiating privately. I include this subsequent analysis in an effort to apply my finding from the previous section to another industry with a slightly different structure. In section 3.7, I suggest that players and teams are learning at the same rate. There is no indication that either side is gaining a systematic advantage through the learning process.

The WERC data that I have provides a binary indicator of the winners of arbitrated cases disputing multidimensional terms, as well as macroeconomic data. I can organize these data as a panel by city or government unit. Using an indicator for hearing wins as the dependent variable allows for an analysis of a model similar to equation 3.4. However, since a city can participate in multiple hearings per year, I do not use a probit analysis. Instead, I use the percentage of cases won by unions against a city as the dependent variable. I regress this percentage on the number of cases the city participated in that year and a trend

variable. I also include the yearly unemployment rate in Wisconsin and the party affiliation of the Wisconsin governor to control for the macroeconomic and political environment. The model appears as follows:

$$\begin{aligned} \text{Percentage of Union Wins} = & \beta_0 + \beta_1 \text{Cases} + \beta_2 \text{Trend} \\ & + \beta_3 \text{Unemployment Rate} + \beta_4 \text{Governor} + \epsilon \end{aligned} \quad (3.5)$$

Table 3.5: Fixed Effects Estimator Controlling for Macroeconomic and Political Factors in Public Sector Arbitration Hearing Outcomes

VARIABLES	(1) FE	(2) AB
Percentage of Union Hearing Wins _{t-1}		-0.182 (0.112)
Yearly Cases in City _t	0.0230 (0.0204)	-0.00112 (0.0520)
Trend	-0.00508*** (0.00177)	0.0464 (0.0428)
Unemployment Rate _t	-0.0270*** (0.00732)	-0.0111 (0.0372)
Governor _t	0.0542** (0.0264)	0.184 (0.142)
Constant	0.603*** (0.0663)	-0.217 (0.748)
Observations	2,217	111
R-squared	0.011	
Number of GovtID	684	62

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

“Yearly Cases in City” is a count variable indicating the number of hearings in within each group participated. “Governor” is a binary variable coded as a one if Wisconsin had governor who identified as a Democrat at the time of the arbitration ruling, zero otherwise.

Coefficients for the control variables are significant and intuitive. I expect a Democratic governor to indicate a more union-friendly political environment. The coefficient is positive and significant at the 95 percent level. On the other hand, as the unemployment rate rises, it should be more difficult for unions to persuade an arbitrator that the time is right for any sort of increase in compensation. This coefficient is negative and significant at the 99 percent level. Yearly Cases in City is not significant in the model, indicating that cities neither benefit nor suffer from increased hearing participation within a given year.

The “Trend” variable is significant and negative suggesting that, over time and when hearings occurs, the arbitration environment has become slightly more favorable to cities than to unions. This advantage may come from the fact that cities can gather more information as they negotiate with multiple unions. However the coefficient, while significant, is very small. Further, the R-squared for the model is extremely small. This indicates that the factors included explain very little of the variation in arbitration wins.

Based on the graph above, cities and unions are clearly learning how to avoid arbitration hearings. However, there is little evidence that cities or unions are gaining an advantage in the hearings that do take place. This is very similar to the MLB analysis in section 3.7, and further supports the hypothesis that parties learn about the arbitration process at the same rate, which promotes private negotiation. Reduced uncertainty during the arbitration process can manifest itself in two ways. Either parties are reducing uncertainty at the same, or differing rates. A party that is able to learn faster will enjoy an increased probability of winning arbitration hearings. This party will prefer to trigger the arbitration option as the likelihood of winning is tipped in their favor. By contrast, if parties are reducing uncertainty at the same rate, the number of hearings should decrease as parties prefer to negotiate privately. However, should a hearing occur, parties

will present offers that are equally likely to sway the arbitrator. In both MLB and WERC data, there is no indication that the respective workers or organizations have increased the probability of winning a hearing over time. However, the use of the arbitration has clearly diminished over the same time frame. This suggests that parties are learning about the process in such a way that private negotiation is promoted without giving an advantage to either party.

3.9 Discussion and Conclusion

MLB data show that teams learn from the arbitration process. The number of arbitration hearings are not only shrinking over time, but the magnitude of impasse is declining. This analysis builds on past literature by modeling the learning process over time, and results largely confirm past findings that MLB's use of arbitration as an impasse resolution mechanism is successful in its role to promote private negotiation.

Uncertainty is an essential element to an arbitration system. MLB data indicate that recent and repeated use reduces uncertainty, but this effect appears to take place slowly. Over the course of 40 years, MLB teams have learned how more efficiently formulate arbitration terms to better avoid the costs of the FOA system. This has allowed promoted private negotiation and allowed parties to gain more control over arbitrated settlements.

This paper explores one of arbitration's two goals. An effect arbitration system should not only promote private bargaining, but hearings, when used, should produce results that resemble privately negotiated settlements. This analysis indicates that MLB's FOA system promotes private bargaining. Further analysis is necessary to indicate whether the process yields results that are similar to privately negotiated settlements (see chapter 2).

BIBLIOGRAPHY

Abrams, R. I. (1999). Before the flood: the history of baseball's antitrust exemption, *Marquette Sports Law Journal* 307

Abrams, R. I. (1999). Inside Baseball's Salary Arbitration Process, *The University of Chicago Law School Roundtable: Vol. 6 (1)*, Article 6.

Akerlof, G. A., & Yellen, J. L. (1990). The fair wage-effort hypothesis and unemployment. *The Quarterly Journal of Economics*, 105(2), 255-283.

Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Annala, C. N., & Winfree, J. (2011). Salary distribution and team performance in Major League Baseball. *Sport Management Review*, 14(2), 167-175.

Ashenfelter, O. C., & Bloom, D. E. (1983). Models of arbitrator behavior: Theory and evidence.

Ashenfelter, O. (1987) Arbitration and the Negotiation Process: Arbitrator Behavior. The AER. Vol 77. No. 2, Papers and proceedings of the ninety-ninth annual meeting of the American Economic Association. pp. 342-346

Babcock, L. C., & Taylor, L. J. (1996). The role of arbitrator uncertainty in negotiation impasses. *Industrial Relations: A Journal of Economy and Society*, 35(4), 604-610.

Barros, C. P., & Garcia-del-Barrio, P. (2008). Efficiency measurement of the English football Premier League with a random frontier model. *Economic modelling*, 25(5), 994-1002.

Baumer, B. S., Jensen, S. T., & Matthews, G. J. (2015). openWAR: An open source system for evaluating overall player performance in major league baseball. *Journal of Quantitative Analysis in Sports*, 11(2), 69-84.

Bazerman, M. H., & Farber, H. S. (1985). Arbitrator Decision Making: when are final offers important?. *Industrial & Labor Relations Review*, 39(1), 76-89.

Bloom, D. E. (1981). Is arbitration really compatible with bargaining?. *Industrial Relations: A Journal of Economy and Society*, 20(3), 233-244.

Bloom, D. E., & Cavanagh, C. L. (1987). Negotiator behavior under arbitration. The AER. Vol 77. No. 2, Papers and proceedings of the ninety-ninth annual meeting of the American Economic Association. pp. 353-358

Bloom, M. (1999). The performance effects of pay dispersion on individuals and organizations. *Academy of Management Journal*, 42(1), 25-40.

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.

Burgess, P. L., & Marburger, D. R. (1993). Do negotiated and arbitrated salaries differ under final-offer arbitration?. *ILR Review*, 46(3), 548-559.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3), 414-427.

Cassing, J., & Douglas, R. W. (1980). Implications of the auction mechanism in baseball's free agent draft. *Southern Economic Journal*, 110-121.

Coleman, B., Jennings, K. M., & McLaughlin, F. S. (1993). Convergence or Divergence in Final-Offer Arbitration in Professional Baseball. *Industrial Relations: A Journal of Economy and Society*, 32(2), 238-247.

Dawson, P., Dobson, S., & Gerrard, B. (2000). Estimating coaching efficiency in professional team sports: Evidence from English association football. *Scottish Journal of Political Economy*, 47(4), 399-421.

DeBrock, L., Hendricks, W., & Koenker, R. (2004). Pay and Performance The Impact of Salary Distribution on Firm-Level Outcomes in Baseball. *Journal of Sports Economics*, 5(3), 243-261.

Deck, C. A., & Farmer, A. (2007). Bargaining over an uncertain value: arbitration mechanisms compared. *Journal of Law, Economics, and Organization*, 23(3), 547-579.

Depken, C. A. (2000). Wage disparity and team productivity: evidence from major league baseball. *Economics letters*, 67(1), 87-92.

Farber, H. S. (1981). Splitting-the-difference in interest arbitration. *ILR Review*, 35(1), 70-77.

Farber, H. S., & Katz, H. C. (1979). Interest arbitration, outcomes, and the incentive to bargain. *ILR Review*, 33(1), 55-63.

Farber, H. S., & Bazerman, M. (1984). The general basis of arbitrator behavior: An empirical analysis of conventional and final-offer arbitration.

Faurot, D. J., & McAllister, S. (1992). Salary arbitration and pre-arbitration negotiation in major league baseball. *ILR Review*, 45(4), 697-710.

Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, 77(3), 953-973.

Franck, E., & Nüesch, S. (2011). The effect of wage dispersion on team outcome and the way team outcome is produced. *Applied Economics*, 43(23), 3037-3049.

Frederick, D. M., Kaempfer, W. H., Ross, M. T., & Wobbekind, R. L. (1998). Arbitration versus negotiation: the risk aversion of players. *Applied Economics Letters*, 5(3), 187-190.

Hofler, R. A., & Payne, J. E. (2006). Efficiency in the National Basketball Association: a stochastic frontier approach with panel data. *Managerial and Decision Economics*, 27(4), 279-285.

Idson, T. L., & Kahane, L. H. (2000). Team effects on compensation: an application to salary determination in the National Hockey League. *Economic Inquiry*, 38(2), 345-357.

Kahn, L. M. (1993). Free agency, long-term contracts and compensation in Major League Baseball: Estimates from panel data. *The Review of Economics and Statistics*, 157-164.

Kamada, T., & Katayama, H. (2014). Team performance and within-team salary disparity: an analysis of nippon professional baseball. *Economics Bulletin*, 34(1), 144-151.

Kennedy, P. (2003). *A guide to econometrics*. MIT press.

Krautmann, A. C., Von Allmen, P., & Berri, D. (2009). The underpayment of restricted players in North American sports leagues. *International Journal of Sport Finance*, 4(3), 161.

Lambrinos, J., & Ashman, T. D. (2007). Salary Determination in the National Hockey League Is Arbitration Efficient?. *Journal of Sports Economics*, 8(2), 192-201.

Lazear, E. P., & Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of political Economy*, 89(5), 841-864.

Leeds, M. A., & Kowalewski, S. (2001). Winner take all in the NFL the effect of the salary cap and free agency on the compensation of skill position players. *Journal of Sports Economics*, 2(3), 244-256.

Levine, D. I. (1991). Cohesiveness, productivity, and wage dispersion. *Journal of Economic Behavior & Organization*, 15(2), 237-255.

Macleon, J. C., Webber, D. A., & Marti, J. (2014). An application of unconditional quantile regression to cigarette taxes. *Journal of Policy Analysis and management*, 33(1), 188-210.

Marburger, D. R. (2004). Arbitrator compromise in final offer arbitration: evidence from Major League Baseball. *Economic Inquiry*, 42(1), 60-68.

Marburger, D. R., & Burgess, P. L. (2004). Can prior offers and arbitration outcomes be used to predict the winners of subsequent final-offer arbitration cases?. *Southern Economic Journal*, 93-102.

Maxcy, J. G. (2004a). Contract length as risk management when labor is not homogeneous. *Labour*, 18(2), 177-189.

Maxcy, J. (2004b). Motivating long-term employment contracts: risk management in major league baseball. *Managerial and Decision Economics*, 25(2), 109-120.

Menger, A. (2015). CLUSTSE: Stata module to estimate the statistical significance of parameters when the data is clustered with a small number of clusters. *Statistical Software Components*.

Miller, M. (1991). *A whole different ball game: The sport and business of baseball*. Birch Lane Press.

Miller, P. A. (2000). A theoretical and empirical comparison of free agent and arbitration-eligible salaries negotiated in major league baseball. *Southern Economic Journal*, 87-104.

Mondello, M., & Maxcy, J. (2009). The impact of salary dispersion and performance bonuses in NFL organizations. *Management Decision*, 47(1), 110-123.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, 1417-1426.

Pappas, D. (2003) "30 Years of Salary Arbitration" *Boston Baseball*

Ramaswamy, R., & Rowthorn, R. E. (1991). Efficiency wages and wage dispersion. *Economica*, 501-514.

Rodriguez-Caro, A., Vallejo-Torres, L., & Lopez-Valcarcel, B. (2016). Unconditional quantile regressions to determine the social gradient of obesity in Spain 1993–2014. *International journal for equity in health*, 15(1), 175.

Roodman, D. (2006). How to do xtabond2: An introduction to difference and system GMM in Stata. *Center for Global Development. Working Paper Number 103*

Samuelson, W. F. (1991). Final-offer arbitration under incomplete information. *Management science*, 37(10), 1234-1247.

Scully, G. W. (1974). Pay and performance in major league baseball. *The American Economic Review*, 64(6), 915-930.

Shaw, J. D. (2014). Pay dispersion. *Annu. Rev. Organ. Psychol. Organ. Behav.*, 1(1), 521-544.

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

Stevens, C. M. (1966). Is compulsory arbitration compatible with bargaining?. *Industrial Relations: A Journal of Economy and Society*, 5(2), 38-52.

Thaler, R. (2012). *The winner's curse: Paradoxes and anomalies of economic life*. Simon and Schuster.

Trevor, C. O., Reilly, G., & Gerhart, B. (2012). Reconsidering pay dispersion's effect on the performance of interdependent work: Reconciling sorting and pay inequality. *Academy of Management Journal*, 55(3), 585-610.

Wittman, D. (1986). Final-offer arbitration. *Management Science*, 32(12), 1551-1561.

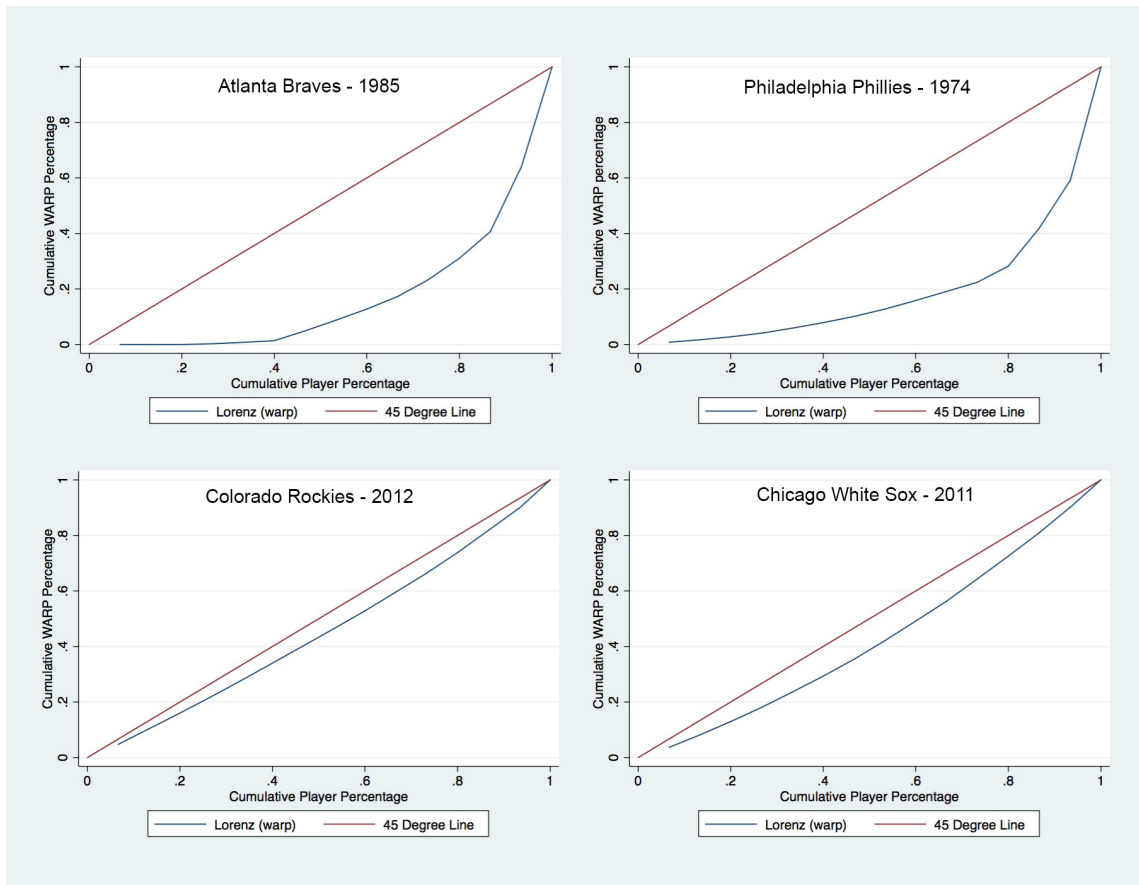
Zimbalist, A. S. (1992). *Baseball and billions: A probing look inside the big business of our national pastime*. Basic Books.

APPENDIX A

CHAPTER 1-LORENZ CURVES

Figure 1 shows Lorenz curves for the 1985 Atlanta Braves, the 1974 Philadelphia Phillies, the 2012 Colorado Rockies and the 2011 Chicago White Sox; the teams with the two highest and two lowest WARP concentrations, respectively.

Figure A.1: Lorenz Curve Comparison



APPENDIX B

CHAPTER 1-ROBUSTNESS

CHECK

Table B.1, on the next page, contains the results of a robustness check in which I use Gini Coefficients as an alternative measure of concentration. To keep the data consistent, I again use the top 15 WARP contributors and top 15 salaries. In order to determine the Gini coefficients for team WARP I first considering the cumulative distribution of individual inputs if each player contributed the exact same level of individual production. Next the actual cumulative distribution of is considered. Finally the Gini Coefficient is calculated by taking the difference between the areas of the perfectly equitable and the actual cumulative distributions, and dividing by the area of the perfectly equitable distribution. Salary Gini coefficients are calculated by the same method. Results are robust to this alternative measure. All signs and levels of significance are consistent with the results of the previous models that use HHI to measure concentration. Additionally, the coefficient on SalaryGini is much larger in the model that is missing WARP Gini. The claim that a model without a direct measure of individual input concentration will suffer from omitted variable bias is robust to this alternative measure of concentration.

Table B.1: Pooled, Fixed Effects, and Arellano-Bond Estimation with and without Salary Measures - Robustness Check

VARIABLES	Dependent Variable is single season Wins, W_t							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	RE	FE	AB	OLS	RE	FE	AB
W_{t-1}				0.0457 (0.0378)				0.217*** (0.0757)
$TeamWARP_t$	1.268*** (0.0390)	1.268*** (0.0464)	1.268*** (0.0456)	1.247*** (0.0452)				
$WARPGini_t$	-12.96*** (4.361)	-13.97*** (3.465)	-14.81*** (3.416)	-15.19*** (4.197)				
$TotalSalary_t^\dagger$.00959 (.0111)	0116 (.0135)	0140 (.0135)	00393 (.0123)	.111*** (.0189)	.101*** (.0353)	.0946*** (.0353)	.0421 (.0289)
$SalaryGini_t$	-9.835*** (2.837)	-11.12*** (3.076)	-12.44*** (3.373)	-8.628** (3.923)	-30.35*** (4.965)	-28.18*** (6.169)	-25.13*** (6.669)	-11.64 (7.186)
Strike1	-13.84*** (1.053)	-13.79*** (1.083)	-13.74*** (1.031)	-14.11*** (1.077)	-22.47*** (1.488)	-22.60*** (1.342)	-22.79*** (1.363)	-23.59*** (1.337)
Strike2	-4.237*** (1.028)	-4.132*** (1.212)	-4.024*** (1.197)	-3.470*** (1.340)	-5.522*** (1.923)	-5.799*** (2.051)	-6.197*** (2.077)	-2.744 (2.144)
Constant	47.75*** (2.265)	48.45*** (2.041)	49.08*** (1.899)	45.30*** (4.125)	87.13*** (1.808)	86.56*** (2.165)	85.96*** (2.338)	66.01*** (7.161)
Observations	558	558	558	524	558	558	558	524
R-squared	0.747		0.741		0.283		0.275	
Number of Teams		30	30	30		30	30	30

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

† in millions