REDUCING ASYMMETRIC INFORMATION: EMPIRICAL EVIDENCE FROM MAJOR LEAGUE BASEBALL

by

Austin G. Hamilton

A thesis submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Master of Science in
Economics

Charlotte

2017

Approved by:

____

Dr. Craig A. Depken, II

____

Dr. Arthur Zillante

____

Dr. Louis Ted Amato

ABSTRACT

AUSTIN G. HAMILTON. Reducing Asymmetric Information: Empirical Evidence from Major League Baseball. (Under the direction of DR. CRAIG A. DEPKEN, II)

This paper estimates the effect wage transparency had on Major League Baseball pitcher's salaries in the mid to late 1980s. The paper implements a conventional method of wage estimation as previously used in the literature. The data, in part are from a new dataset of players' salaries, which includes salaries from the pre-1985 period. The results show that on average, Major League Baseball pitchers' salaries increase by 48.3% over the 1984 level after transparency. However, in 1986 the salaries decline to 30.3% over the 1984 level, a 18% decline compared to 1985. The paper discusses some possible reasons why there was a decline in 1986. One such explanation is that the momentum of upward pressure on wages made the average wage overshoot the equilibrium wage by 18%. Conversely, the period of owner-to-owner collusion is believed to have taken place during 1985, 1986, and the 1987 seasons, which likely put downward pressure on wages during that period. To adjust for collusion, the paper interpolates damages awarded by the courts in 1990 for collusion and estimates the collusion adjusted transparency effect to be 51.3%

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: Introduction

Let us consider a thought experiment: you are interested in buying a car, say a 2017 Mustang GT, which seems like great value for a sports car of its caliber. You decide you cannot live without it and, like almost every other person wanting to buy a car, you desire to pay the lowest amount a dealer is willing to charge. However, in the real world it is nearly impossible to know that exact value, so like any rational and methodical buyer, you estimate that value before you make the purchase. This spread between what is paid and the lowest amount the dealer is willing to sell the car for is one of the predominant metrics of salesmanship. The salesperson is usually the only participant in the transaction who knows this spread. However, the salesperson has inverse incentives compared to the buyer; his or her commission is directly proportional to how well he or she exploits this asymmetric information in negotiations. A 2017 Mustang GT's MSRP is $33,195[1]; however, services like TRUEcar claim you can get it as low as $30,969. What would happen if you knew the absolute lowest price for which a dealership would sell a given Mustang GT? You would likely not pay any more than that lowest price.

To further this thought experiment: let us now imagine Congress passing legislation that requires all Ford dealerships to disclose the lowest price each would be willing to take for a Mustang GT. Assuming they do not respond with a counter strategy, the price at which you would pay for a Mustang GT would converge to that lowest price. Instead, now imagine transparency was applied to the other side of the transaction, where the salesperson knows your maximum willingness to pay. The salesperson would likely sell the Mustang at the highest price, which would maximize their potential

[1] https://www.truecar.com/prices-new/ford/mustang-pricing/2017/077B5BBE/

commission. Consequently, transparency creates a price change in favor of the party receiving the asymmetric information.

The previous case seems rudimentary when talking about the prices of goods, but does this principle of transparency hold when you apply it to labor markets? If you were to inform employees at a given firm what other employees made in similar positions, would there be a convergence in wages? In the literature, there has been some research on transparency applied to executive compensation for publically traded corporations in Europe (Schmidt 2012), which similarly found that transparency had an upward effect on compensation. Thus, hypothesizing transparency might have transcendent property within wage markets. When transparency is applied to employees' wages, they move upward. Transparency might even have the potential to become a future tool of policy makers. The question this paper addresses is does transparency have a similar effect for Major League Baseball pitchers.

Many informed people within Major League baseball believed during the early 1980s that baseball salaries did not represent a fair wage when compared to players' marginal product. People like Marvin Miller, then executive director of the Major Leagues Baseball Players Association, were outspoken proponents for wage transparency within Major League Baseball. He believed that managers and owners were exploiting players during this period. The pressure from players, and other members of the baseball community in the early 1980s, culminated in 1985 when every Major League Baseball player found out what other players were earning. This leads us to question: did this new transparency change baseball players' earnings? To many others and my own surprise, this question could not be answered until recently as there was not a readily available database of players' salaries from the pre-1985 period.

CHAPTER 2: Literature Review

For decades, many labor economists have looked at Major League Baseball for its rich and unique data. Unlike many other industries, a player's marginal revenue product is more easily defined by their ability to contribute to winning a baseball game, which then increases attendance, thus increasing revenues. The literature widely asserts that player salaries are primarily a function of their expected performance. In practice, managers and owners estimate expected performance by the respective player's past performance, injury history, and years played in Major League Baseball. Some additional explanatory variables of a player's salary are a function of a given team's revenues. Yet, due to data limitations researchers must estimate or proxy this variable by estimating revenues through taking the product of average ticket price and number of attendees, plus estimates for concession, memorabilia, and media revenues for a given team. Alternatively, some researches proxy these revenues by using population and average income data for a team's geography (Stone and Pantuosco 2008 [1]).

The literature has also identified structural patterns to Major League Baseball players' salaries and years played in the Major League Baseball. This has several causes, first a player is rewarded for more experience, which is typically proxied by years playing in Major League Baseball. In addition, since 1973 after two years in Major League Baseball a player may enter arbitration, and since 1976 after six years a player may become a free-agent. The impacts of arbitration and free agency are difficult to disentangle from experience alone. However, the literature has consistently found that both arbitration and free agency lead to higher salaries, but free agency also increased contract duration (Kahn 1993 [2]).

Some of the foundational work in identifying the salary question for Major League Baseball and estimation of marginal revenue product was offered by Scully (1974). This paper still shapes how many researchers estimate Major League Baseball players' salaries. Scully divides players into two primary categories: hitters and pitchers. Since this paper only estimates the wages of pitchers, only the independent variables, which are typically included in a pitcher's wage equation are discussed. Scully first identified that the earned run average, what was thought as the primary metric of a pitcher's performance, was virtually uncorrelated with his salary. Scully argues the single-best metric for explaining a pitcher's performance is his strikeout-to-walk ratio. In addition to the strikeout-to-walk ratio, Scully also included the number of years in the majors, the ratio of innings pitched to total innings, the population of the team's market, a team specific estimate of marginal revenue, and a dummy variable for playing in the National League. Summers and Quinton (1982 [3]) extended Scully's work and analyzed the effects free agency had on his previous wage estimation model. They show that free agency did increase player's wages. The next major contribution in the Major League Baseball player salary estimation was Kahn (1993), who affirmed that free agency did increase player's salaries, but that free agency also encouraged longer contracts right before free agency eligibility. Unlike previous models, Kahn also shows that arbitration increases player's salaries. Kahn's pitcher-specific explanatory variables are career winning percentage (WPCT), career earned run average (ERA), percentage of games started (PS), percentage of appearances resulting in saves (PSV), interaction between PS and WPCT, interaction between PS and PSV, and interaction between PS and ERA.

Krautmann and Oppenheimer (2002 [4]) show that contract length has an impact on both the player's wage equation and performance outcomes. They find that players or owners make a tradeoff between contract length and total salary, because of the player's and owner's risk preferences (Maxcy 2004 [5]). Another confounding feature to longer contracts is that players appear to perform worse after signing longer contracts. Krautmann, Gustafson, and Hadley (2003 [6]) find that whether a player is a starter, stopper or long-reliever changes the wage equation and that a separate equation should be estimated for each pitcher type.

Other studies that looked at the impact of the dissolution of this type of asymmetric information on wage markets found examples with corporate executives and city managers. As previously mentioned in the introduction, Schmidt (2012 [7]) looked at German companies traded on the DAX from the late 2000s and concluded that more compensation disclosure lead to higher compensation. Conversely, Mas (2014 [8]) evaluated a 2010 policy that California cities had to disclose what their city managers were being paid. However, in this case Mas observed approximately an 8% decline in salaries after transparency.

CHAPTER 3: Methodology

The body of research of professional baseball player wages separates players into two broad categories: pitchers and hitters. This is because of the differences in the metrics used to measure the quality of each type of player. Here, I only look at pitchers [1]. The pre-1985 salary data came from salary data collected from the Baseball Hall of Fame by Michael Haupert, currently at the University of Wisconsin - La Crosse. All player statistics and post-1984 salary data are from the Sean Lahman baseball database. Based on the previous literature of baseball salaries, I use common metrics of pitcher quality: the player's start ratio, the team's fan base (attendance), eligibility for arbitration, eligibility for free agency, which arm he pitches with, his strikeout-to-walk ratio, age, and number of outs pitched.

It is frequently thought there are different categories of pitchers; some of the most common are starters, relievers, and closers. For example, starters, usually pitch first and have the ability to pitch for several innings. Teams usually have several starting pitchers, because of the level of harm inflicted on the body after pitching for several innings. Relief pitchers typically takeover for starting pitchers to give them rest or to rescue the starting pitcher if he is not performing well that day. Relief pitchers are considered the most versatile of pitchers in style or technique. Closing pitchers are valued for their consistency and ability to get strikeouts.

A team's fan base has many other highly correlated alternatives, such as population density in the surrounding geographic area, but in this paper for the availability of data attendance is used as a proxy for fan base. As outlined in the Basic Agreement[2],

[1] When isolating the period of interest, 1984 to 1985, the Haupert data had a substantially larger number of pitchers than hitters.

[2] http://www.baseballprospectus.com/compensation/cots/league-info/cba-history/

players can enter arbitration after two years of service in Major League Baseball. In addition, players can become a free agent after six years of service in Major League Baseball. Both arbitration and free agency reduce teams monoponistic power over their players by circumventing the reserve clause.

Today, approximately 25%[3] of Major League Baseball players are left-handed. In contrast, only 10%[4] of the general population of adult males are left-handed. In professional baseball, it is statistically and logically evident that there is a preference for left-handed pitchers. Pitchers are commonly evaluated by their strikeout-to-walk ratio. The reason for this is that it captures a pitcher's ability to generate strikeouts, but then penalizes him for inconsistency. Experience is another player attribute that is valued by teams. A player with more experience adds value by encouraging other players, by word and action. Furthermore, a more seasoned player should be more consistent in play, not being as susceptible to nerves. A major factor influencing a player's salary is the team's budget. A team's budget is highly dependent on the team's fan base. However, as previously noted, fan base is proxied for by attendance.

A player's salary is primarily determined by his ability to generate revenue for his team (Scully 1974). For pitchers, this may be proportional to their ability to prevent runs scored by the other team. Because fans of the team prefer winning over losing, winning games should correlate with more ticket, concession, memorabilia, and media revenues. Still, the most elusive attribute that can contribute to a team's revenues is fame or draw of a particular player. This has been proxied by the number of jersey sales or the value of the player's baseball cards; however, due to data limitations this is not accounted for in the models presented in this paper.

There are many other factors influencing a player's salary. The commonly measured determinants of a player's salary can be separated into three categories: player level, team level, and league level. Player level determinants include the number of outs

[3] http://www.livescience.com/2665-baseball-rigged-lefties.html

[4] http://www.huffingtonpost.com/2012/10/29/left-handed-facts-lefties_n_2005864.html

pitched, pitcher type or age. Another example is the ability of a player or the firm, which he hires, to negotiate his salary. Team level determinants include annual payroll budget for a team, playing strategy, or demand for a particular type of player. Lastly, league level determinants include minimum salary levels, arbitration requirements or collective bargaining restrictions. For example, during the first two years of a Major League Baseball player's career, he is subject to the reserve clause, giving the team monopsony power over the player, which limits wage growth. There are some determinants that can overlap in multiple categories, such as the substitutability of a player, this could be at the level of the player, team, or both.

CHAPTER 4: Data

The period of most interest is from 1984 to 1985, when transparency was applied to Major League Baseball salaries. There were 31 pitchers for whom we have salary data for those two years. It is believed that this sample is a random draw from the population of Major League Baseball pitchers during these years. There were 461 pitchers who played in either the 1984 or 1985 seasons, so the 31 players make up approximately 6.7% of all pitchers from that period. Comparing only pitchers who played in both 1984 and 1985 seasons, the number falls to 283 and these 31 players comprise approximately 11.0% of all pitchers that played in both of those years. The table and graphs below list the 31 players from the sample.

Two players experience declines in their nominal salaries in 1985 compared to 1984: Britt Burns and Bob Stoddard. Four players experience declines in their nominal salaries in 1986 compared to 1984: Doug Bair, Bill Campbell, Sammy Stewart, and Bob Stoddard.

The following section provides a graph of each of the 31 pitchers nominal salaries. In regard to the graphs below: green triangles represent the year a player signed as a free agent; orange triangles represent the year a player signed as free agent, but we do not have salary data for those years; a brown square represents the year a player debuted; a magenta diamond represents the seventh year after a player debuted, that is, he became eligible for free agency; a gold diamond represents the year free agency began.

There are two datasets used to estimate the model, the first, which was previous mentioned, had the criteria that a pitcher played in both the 1984 and 1985 seasons, but due to data limitations, the number of total pitchers were only 31. However, the

Table 4.1: List of Players (Small Sample)

Player	**First Year**	**Last Year**	**Span**
Don Aase	1977	1989	12
Doyle Alexander	1971	1989	18
Keith Atherton	1983	1988	5
Doug Bair	1976	1990	14
Salome Barojas	1983	1988	5
Joe Beckwith	1979	1986	7
Steve Bedrosian	1981	1995	14
Rick Behenna	1983	1985	2
Mike Bielecki	1984	1997	13
Greg Booker	1983	1990	7
Rich Bordi	1980	1986	6
Warren Brusstar	1977	1985	8
Britt Burns	1978	1985	7
John Butcher	1980	1986	6
Jeff Calhoun	1984	1988	4
Bill Campbell	1974	1987	13
Rick Camp	1976	1985	9
Steve Carlton	1965	1988	23
Don Carman	1983	1991	8
Dennis Eckersley	1975	1998	23
Rollie Fingers	1968	1985	17
Nolan Ryan	1966	1993	27
Tom Seaver	1967	1986	19
Bob Shirley	1977	1986	9
Eric Show	1981	1990	9
Doug Sisk	1982	1988	6
Nate Snell	1984	1987	3
Julio Solano	1983	1988	5
Sammy Stewart	1978	1987	9
Bob Stoddard	1981	1987	6
Tim Stoddard	1975	1988	13

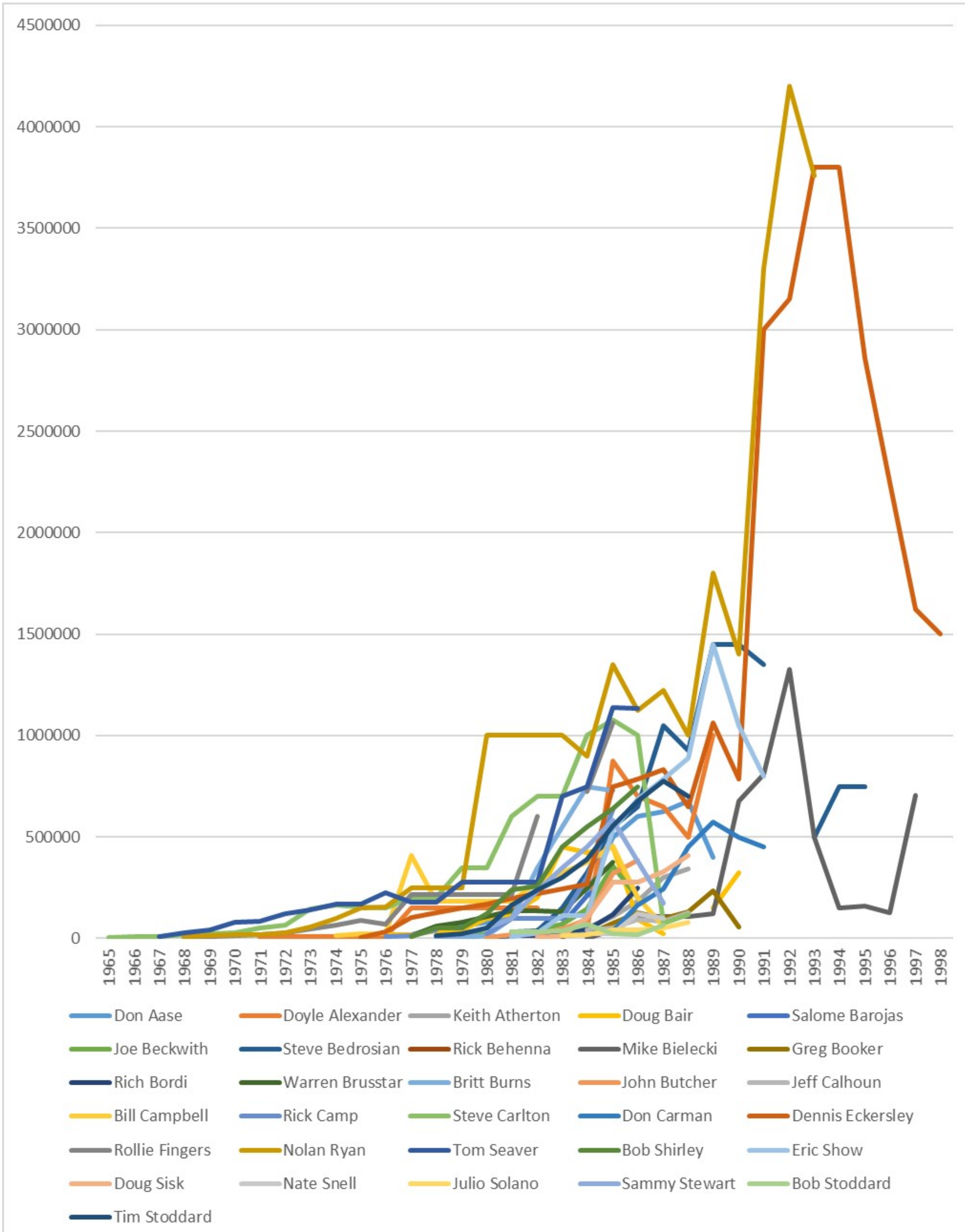

Figure 4.1: Player's Salaries (Nominal)

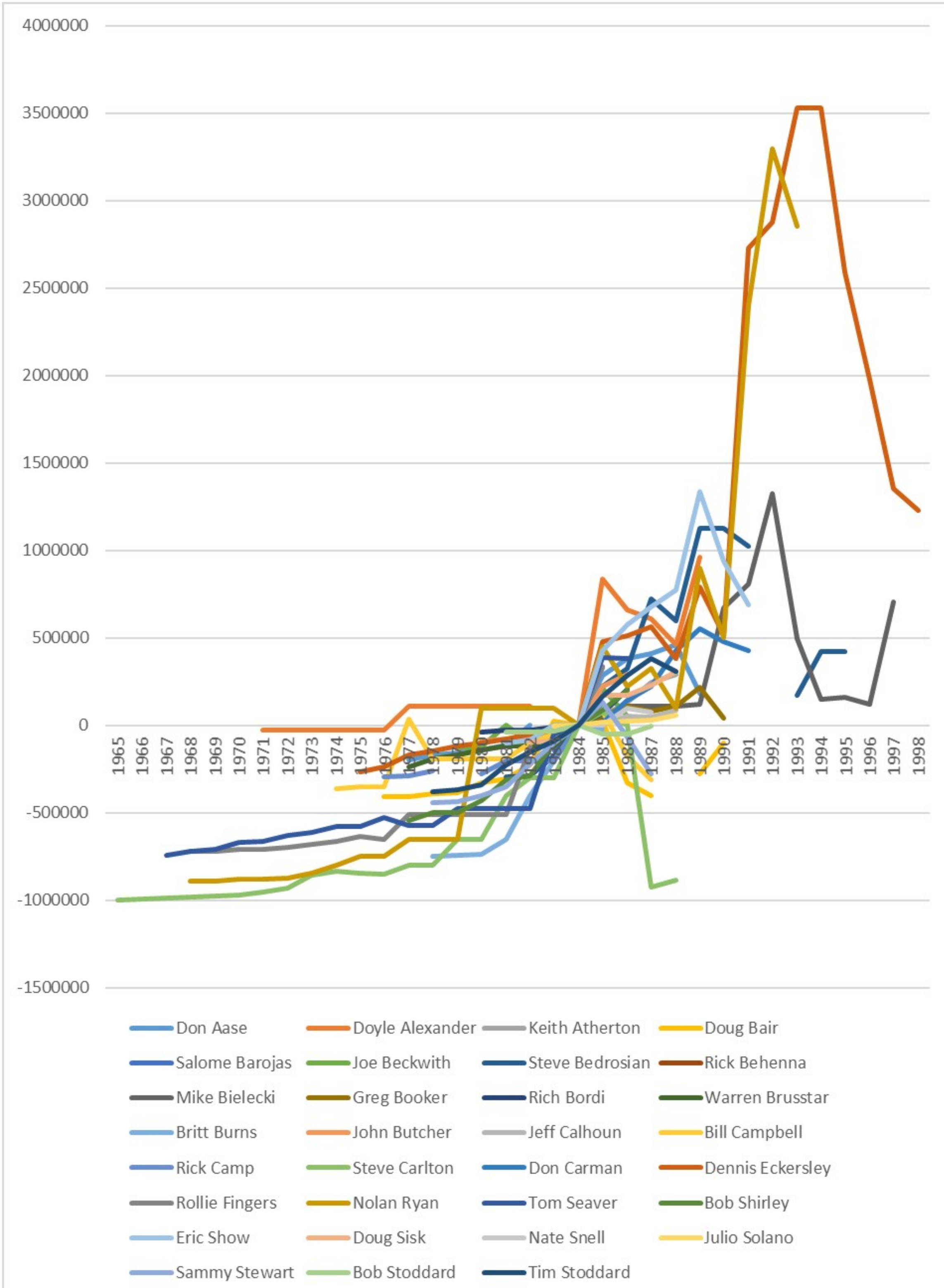

Figure 4.2: Player's Salaries (Nominal and Normalized to 1984)

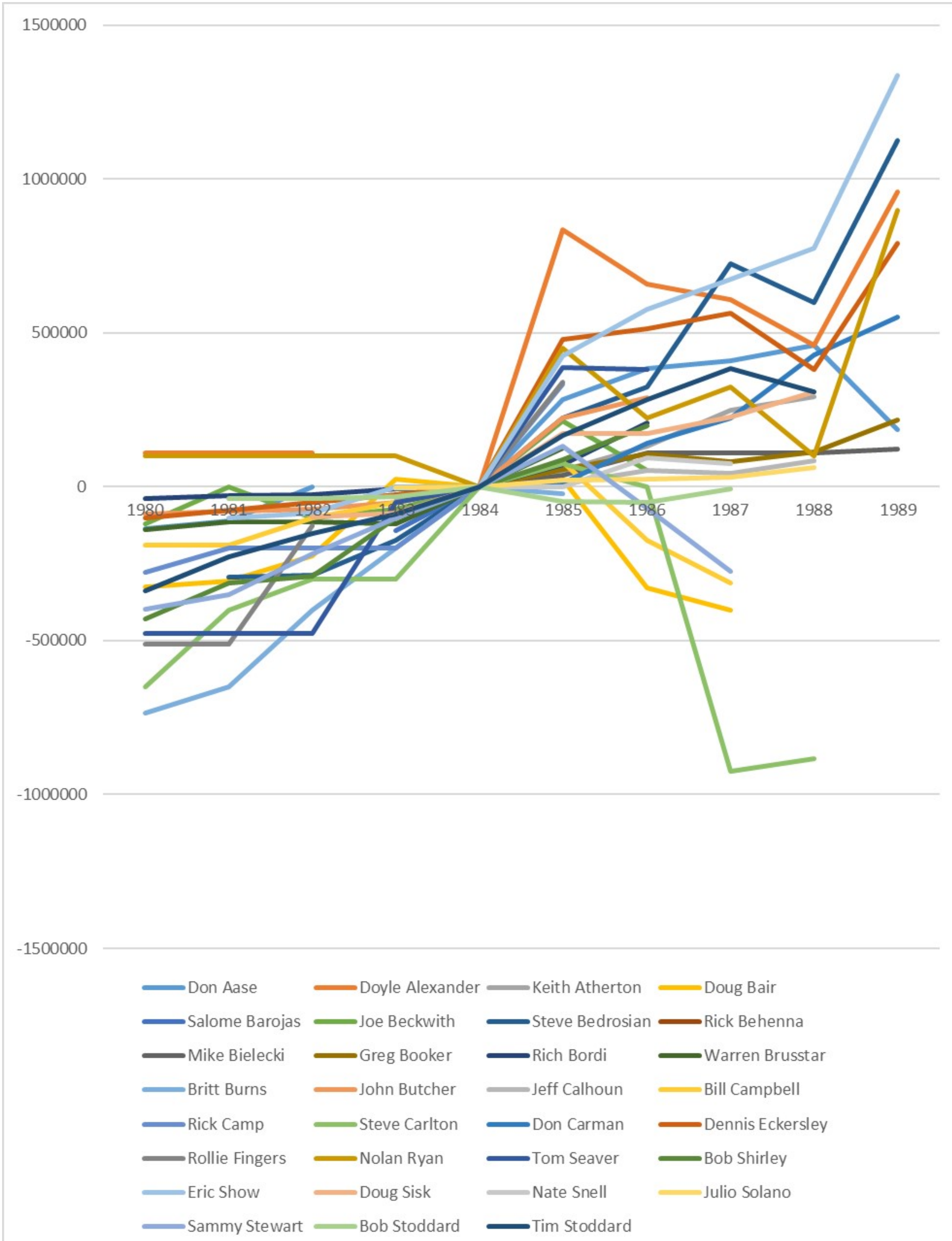

Figure 4.3: Player's Salaries (Nominal and Normalized to 1984) 1979-1990)

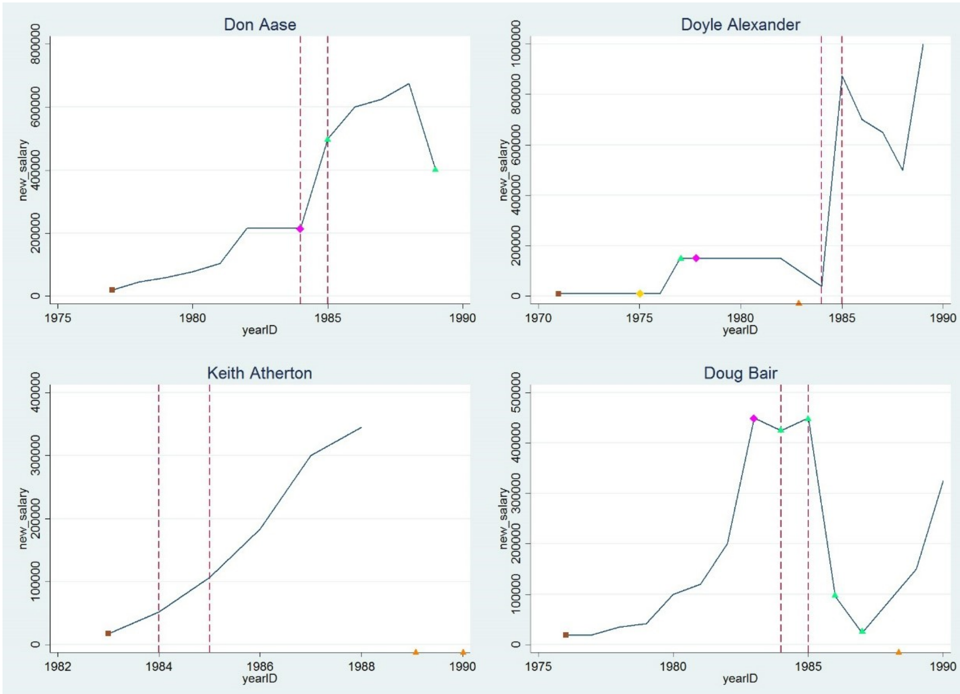

Figure 4.4: Player's Salaries (Nominal) Aase to Bair

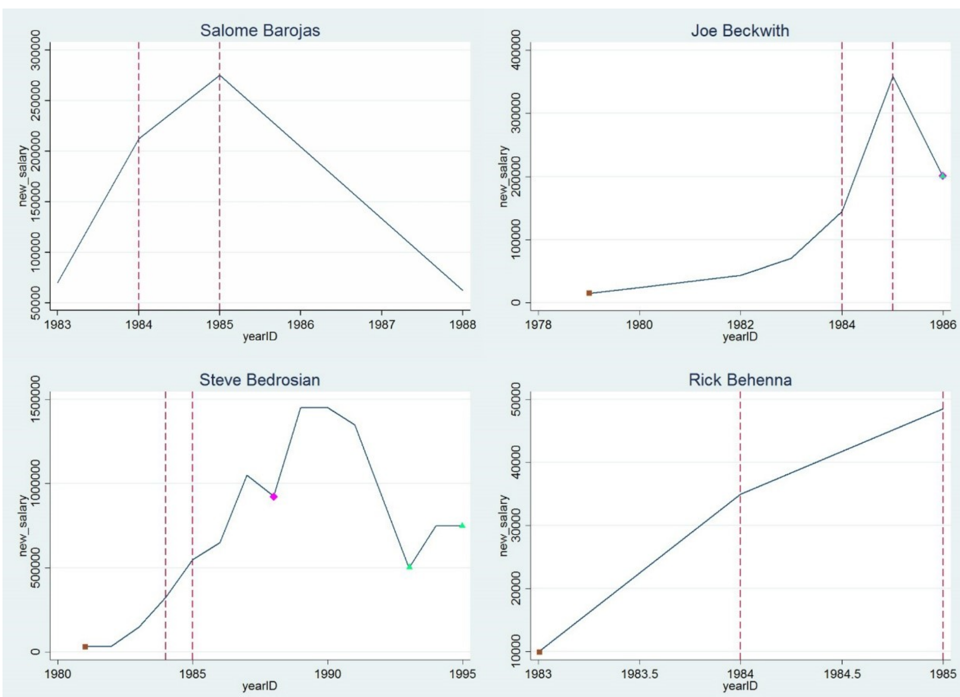

Figure 4.5: Player's Salaries (Nominal) Barojas to Behenna

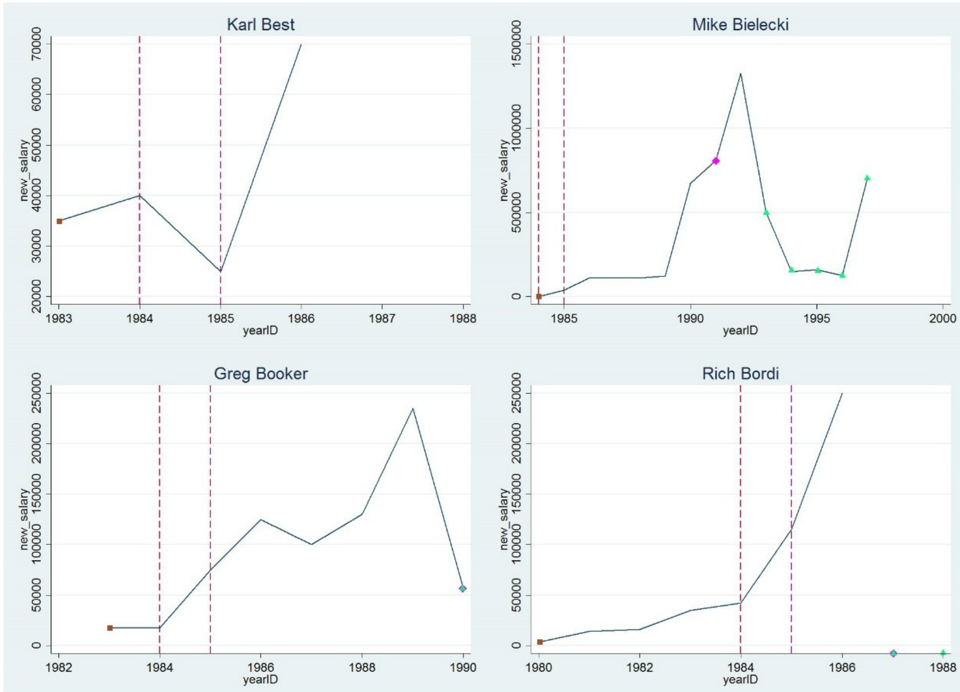

Figure 4.6: Player's Salaries (Nominal) Best to Bordi
1

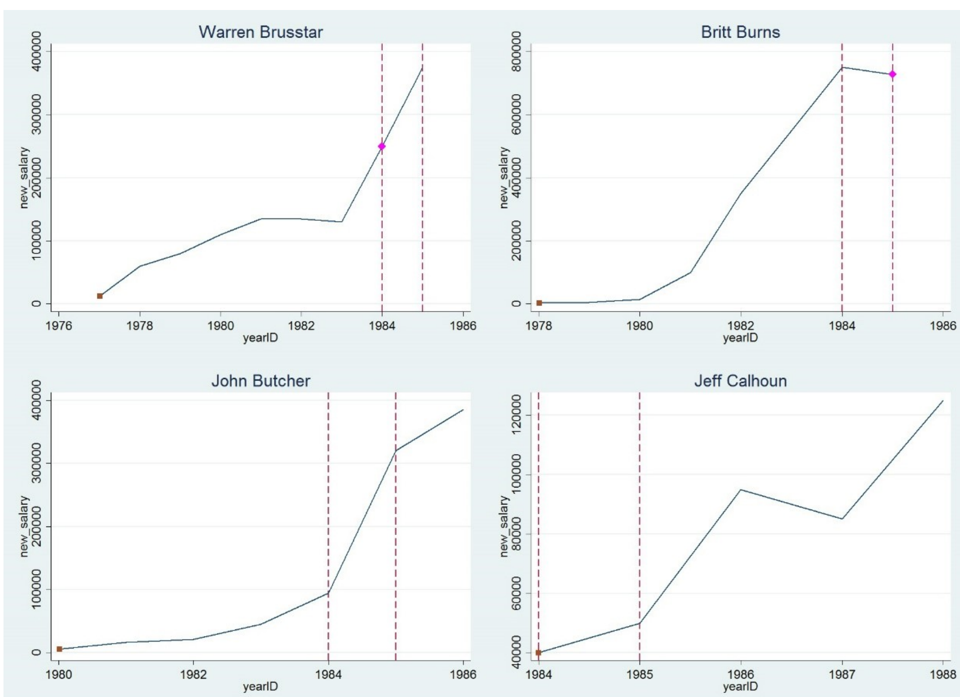

Figure 4.7: Player's Salaries (Nominal) Brusstar to Calhoun

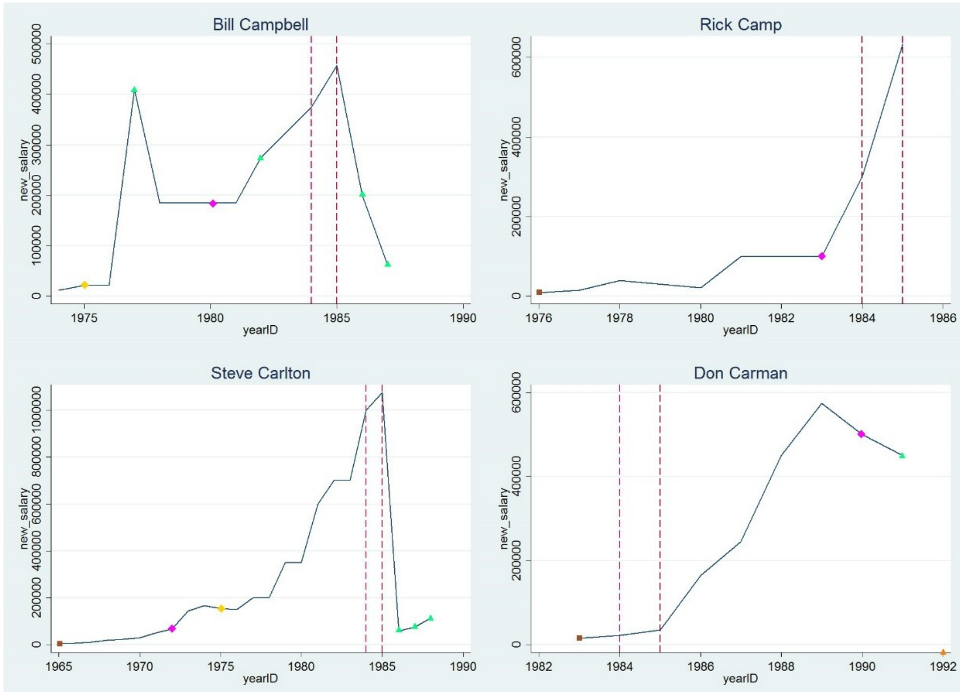

Figure 4.8: Player's Salaries (Nominal) Campbell to Carman

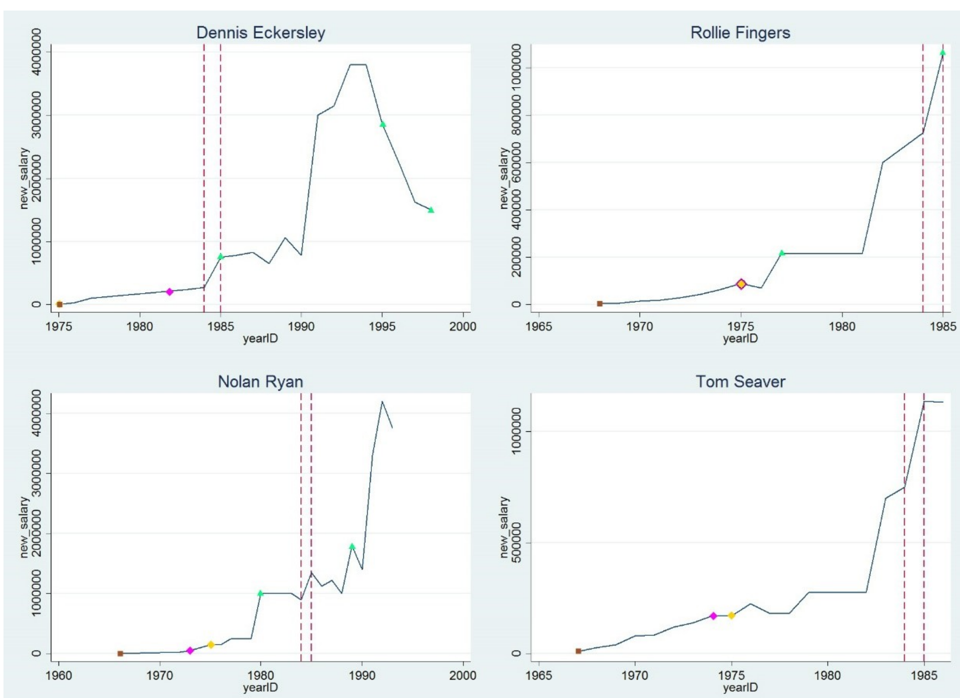

Figure 4.9: Player's Salaries (Nominal) Eckersley to Seaver

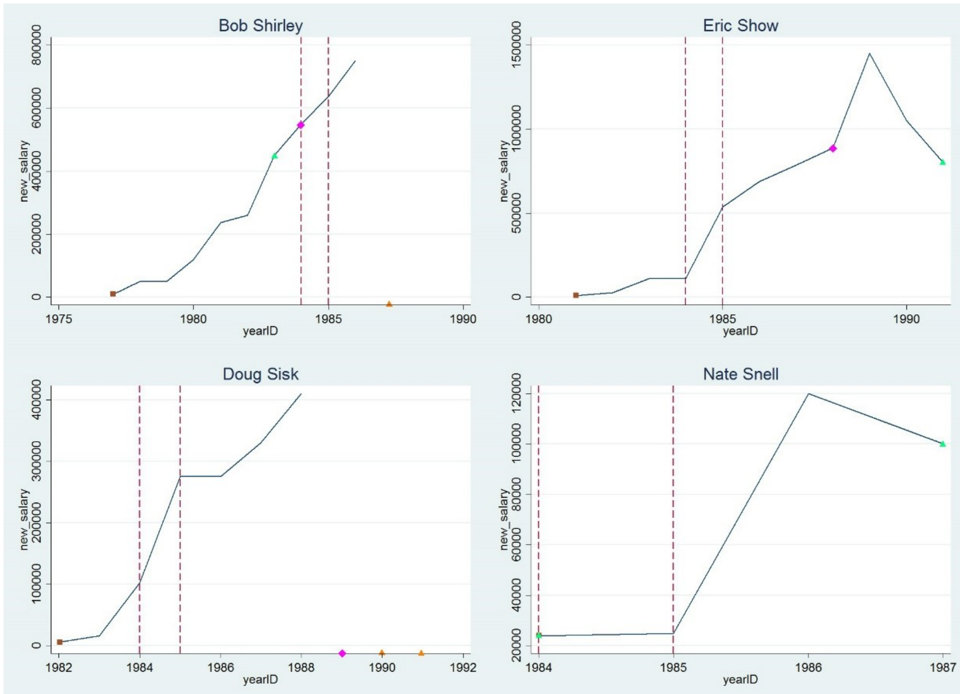

Figure 4.10: Player's Salaries (Nominal) Shirley to Snell

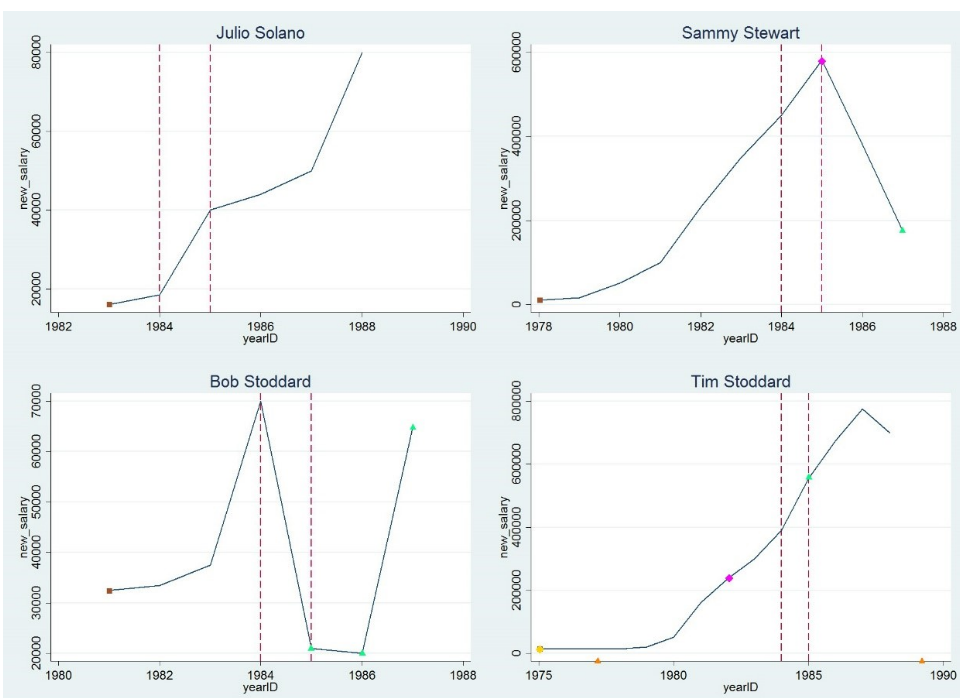

Figure 4.11: Player's Salaries (Nominal) Solano to Stoddard

second sample required only that a pitcher played in either 1984, 1985 or 1986, which increased the total number of pitchers to 334. The larger sample accounts for 65.9% of all pitchers who played in either the 1984, 1985 or the 1986 seasons.

Table 4.2: Summary Statistics: Smaller Sample (Played in both 1984 and 1985)

Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
salary_nom	347	366,894	597,679	1,200	4,200,000
yr_{1984}	358	0.087	0.282	0	1
yr_{1985}	358	0.087	0.282	0	1
yr_{1986}	358	0.070	0.255	0	1
yr_{1987}	358	0.061	0.240	0	1
$post_{1987}$	358	0.187	0.391	0	1
starter	358	0.416	0.494	0	1
stopper	358	0.411	0.493	0	1
inter_closing_starting	356	0.027	0.045	0	0.250
$closing\_ratio_{ma2l}$	356	0.297	0.270	0	0.935
$(closing\_ratio_{ma2l})^2$	356	0.161	0.221	0	0.874
$start\_ratio_{ma2l}$	356	0.399	0.418	0	1
$(start\_ratio_{ma2l})^2$	356	0.333	0.415	0	1
$\ln(IPOuts_{ma2l})$	296	5.688	0.771	1.79	6.90
$\ln(attendance_{ma2l})$	296	14.336	0.331	13.46	15.09
arbitration	358	0.757	0.430	0	1
free_agency	358	0.430	0.496	0	1
left_throw	358	0.162	0.369	0	1
$\ln(strikeout\_to\_walk_{ma2l})$	292	0.633	0.486	-0.882	2.91
age	358	30.025	5.431	19	46
$(age)^2$	358	930.919	344.588	361	2,116

Table 4.3: Summary Statistics: Larger Sample (Played in either 1984, 1985 or 1986)

Variable	Obs	Mean	Std. Dev.	Min	Max
salary_nom	2,270	1,131,653	1,737,362	4,200	18,000,000
yr_{1984}	2,860	0.001	0.032	0	1
yr_{1985}	2,860	0.130	0.337	0	1
yr_{1986}	2,860	0.136	0.342	0	1
yr_{1987}	2,860	0.112	0.316	0	1
$post_{1987}$	2,860	0.574	0.495	0	1
starter	2,860	0.456	0.498	0	1
stopper	2,860	0.291	0.454	0	1
inter_closing_starting	2,860	0.029	0.044	0	0.250
$closing\_ratio_{ma2l}$	2,860	0.236	0.238	0	1
$(closing\_ratio_{ma2l})^2$	2,860	0.112	0.184	0	1
$start\_ratio_{ma2l}$	2,860	0.446	0.411	0	1.00
$(start\_ratio_{ma2l})^2$	2,860	0.368	0.411	0	1
$\ln(IPOuts_{ma2l})$	1,981	5.590	0.788	1.792	6.804
$\ln(attendance_{ma2l})$	1,981	14.487	0.310	13.336	15.213
arbitration	2,860	0.818	0.386	0	1
free_agency	2,860	0.515	0.500	0	1
left_throw	2,860	0.348	0.476	0	1
$\ln(strikeout\_to\_walk_{ma2l})$	1,954	0.609	0.392	-1.098612	2.108429
age	2,860	30.601	4.782	19	50
$(age)^2$	2,860	959.286	304.614	361	2,500

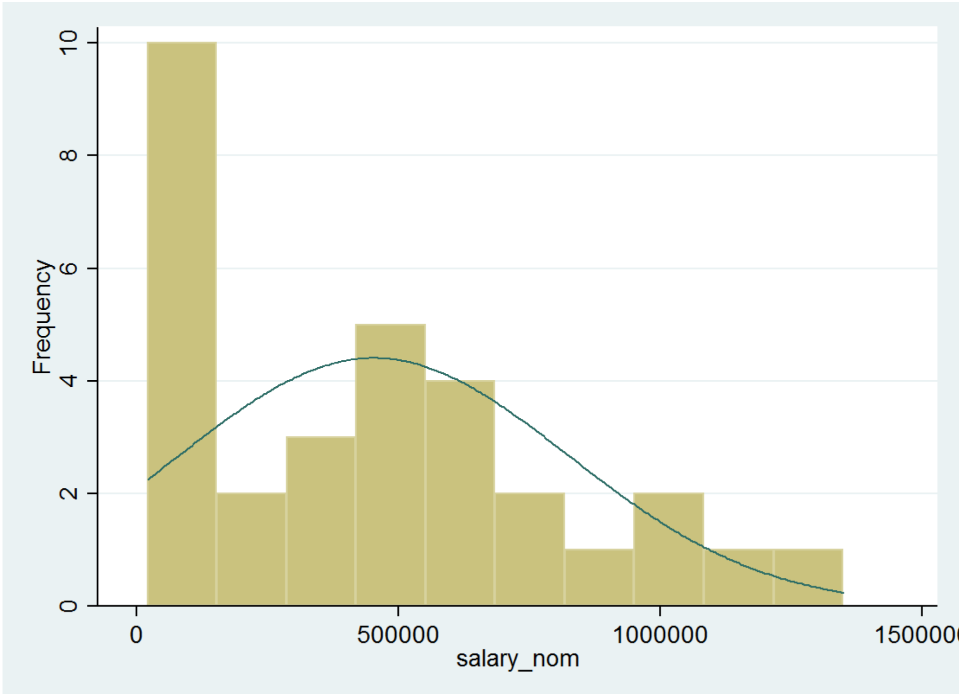

Figure 4.12: Player's Nominal Salaries (Small Sample)

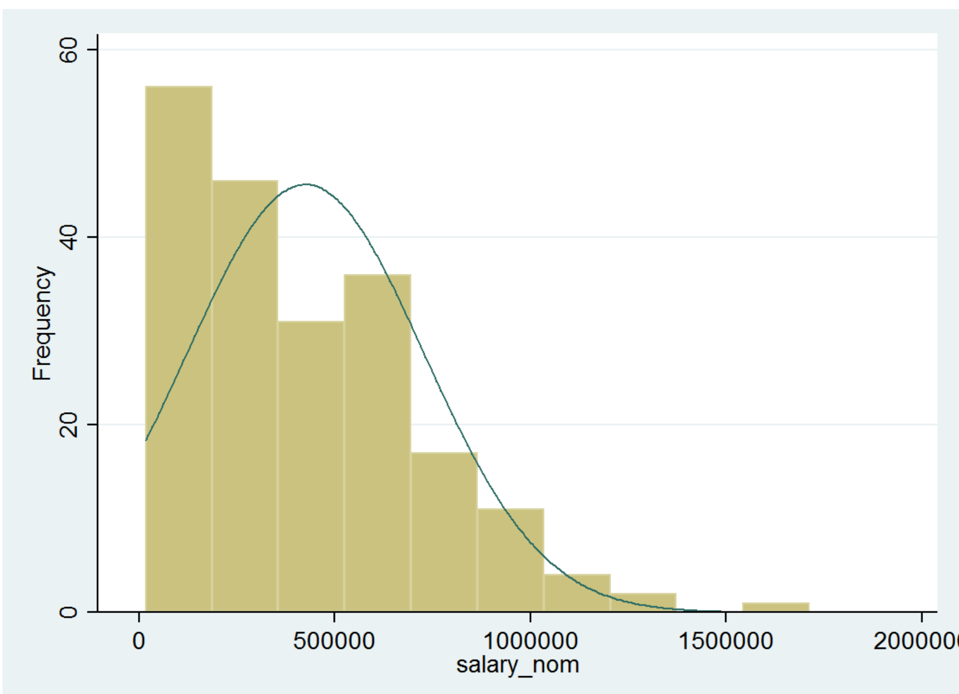

Figure 4.13: Player's Nominal Salaries (Large Sample)

CHAPTER 5: Model

The empirical models includes many of the common metrics used to measure the performance of a pitcher. In addition to metrics for pitcher performance, I control for attendance, which proxies for a given team's fan base. In addition, free agency and arbitration eligibility are included in the model. I also control for if a player is a starter, long-reliever, or closer in one of the three models. For a player's start ratio, outs pitched, attendance, and strikeout-to-walk ratio, I use a moving average from the previous two seasons. The rational for a two-season moving average is that a pitcher's salary is based on the expectation of future performance, which is impossible to predict perfectly. That is why, teams and owners estimate a pitcher's future performance as a function of his past performance. In addition, using moving averages allows for more degrees of freedom than including two lagged dependent variables for each explanatory variable. Instead of moving averages, it is common practice in the literature to use a player's career average. However, a career average weights every season's performance equally, when in the real world the most recent performance is weighted heaviest when determining a player's salary. Lastly, natural logs of outs pitched, attendance, and strikeout-to-walk ratio are used to smooth the data.

As defined by the Basic Agreement, a player is eligible for arbitration after two years on a Major League Baseball team. The arbitration dummy variable takes a value of one if the difference by which a given year and the player's debut year is greater than two. As defined by the new Basic Agreement, a player is eligible for free agency after 6 years on a Major League Baseball team. The free agency eligibility dummy variable takes a value of one if the difference by which a given year and the player's debut year is greater than six, and the given year is after 1974. The start and closing ratio both have non-linear properties; in contrast to creating dummy variables for starters, relievers, and closers, a quadratic was used to capture the difference in the three types of pitchers. In addition, an interaction term for the start and closing ratios is included in the model. Age captures maturity, experience and deterioration of the body. The first two impacts would augment a player's value, the latter would detract from a player's value. Salaries are converted to real prices by multiplying by CPI, and transformed into natural logarithms to accommodate skew (see figure 4.12 4.13 on page 20).

5.1 The Estimation Equation Is Specified As:

Model 1:

$$\ln(\text{salary}_R)_{it} = \beta_0 + \beta_1(yr_{1984})_i + \beta_2(yr_{1985})_i + \beta_3(yr_{1986})_i + \beta_4(yr_{1987})_i + \beta_5(post_{1987})_i$$
$$+\beta_6(start\_ratio_{ma2l})_{it} + \beta_7(start\_ratio_{ma2l})^2{}_{it} + \beta_8(arbitration)_{it} + \beta_9(free\_agency)_{it}$$
$$+ \beta_{10}(left\_throw)_i + \beta_{11}(age)_{it} + \beta_{12}(age)^2{}_{it} + \beta_{13}ln(attendance_{ma2l})_{it} + \varepsilon_{it}$$

Model 2:

$$\ln(\text{salary}_R)_{it} = \beta_0 + \beta_1(yr_{1984})_i + \beta_2(yr_{1985})_i + \beta_3(yr_{1986})_i + \beta_4(yr_{1987})_i + \beta_5(post_{1987})_i$$
$$+ \beta_6(starter)_{it} + \beta_7(stopper)_{it} + \beta_8(arbitration)_{it} + \beta_9(free\_agency)_{it}$$
$$+ \beta_{10}(left\_throw)_i + \beta_{11}(age)_{it} + \beta_{12}(age)^2{}_{it} + \beta_{13}ln(IPOuts_{ma2l})_{it}$$
$$+ \beta_{14}ln(strikeout\_to\_walk_{ma2l})_{it} + \beta_{15}ln(attendance_{ma2l})_{it} + \varepsilon_{it}$$

Model 3:

$$\ln(\text{salary}_R)_{it} = \beta_0 + \beta_1(yr_{1984})_i + \beta_2(yr_{1985})_i + \beta_3(yr_{1986})_i + \beta_4(yr_{1987})_i + \beta_5(post_{1987})_i$$
$$+\beta_6(start\_ratio_{ma2l})_{it} + \beta_7(start\_ratio_{ma2l})^2{}_{it} + \beta_8(arbitration)_{it} + \beta_9(free\_agency)_{it}$$
$$+ \beta_{10}(left\_throw)_i + \beta_{11}(age)_{it} + \beta_{12}(age)^2{}_{it} + \beta_{13}ln(IPOuts_{ma2l})_{it}$$
$$+\beta_{14}ln(strikeout\_to\_walk_{ma2l})_{it} + \beta_{15}(inter\_closing\_starting)_{it} + \beta_{16}(closing\_ratio_{ma2l})_{it}$$
$$+ \beta_{17}(closing\_ratio_{ma2l})^2{}_{it} + \beta_{18}ln(attendance_{ma2l})_{it} + \varepsilon_{it}$$

where ε_{it} is a zero-mean error term, the $\beta's$ are parameters to be estimated, i indexes the player, and t indexes the year.

As mentioned in the introduction, the focus of this paper is the effects of transparency on wages within Major League Baseball. To address this question, five year dummy variables are included in the model, the base period is the period prior to 1984. The reason for this method of separating the years between 1984 and 1987 is to estimate if transparency had any long-term impact on salaries and which years experienced statistically significant real changes in salary levels.

Several variables that were not included in this model, but are frequently included in many papers in the literature. Some of those are wins, losses, earned runs and saves. When including variables such as wins, it is challenging to disentangle the impact that other members of the team had on winning and the individual contribution of the pitcher. Scully (1974) argued that earned runs was virtually uncorrelated with a pitchers salary; therefore, it was omitted from the model.

A random effects panel estimator was used in the analysis, because the group of players studied are believed to be a random sample of the population of pitchers in Major League Baseball. In addition, the random effects estimator offers increased efficiency gains over a fixed effects estimator, which is ideal given the small sample analyzed here. Another benefit to the random effects model is that time-invariant variables such as left-handedness of a player, can be included. However, the random effects estimator is more sensitive to omitted variable bias than a fixed effects model. Furthermore, the random effects model is vulnerable to being inconsistent if player fixed effects are correlated with one or more of the dependent variables. In addition, the standard errors are calculated by clustering of individual players. The model's methodology will be further discussed in the results section of this paper.

CHAPTER 6: Results

The tables below display the random effects estimator's results. The table compares three different model's estimates on both the small and large samples using a random effects estimator.

The comparison period for all year dummy variables is pre-1984. The yr_{1984} dummy variable is not statistically different from zero in any of the models. The yr_{1985} dummy variable is statistically different from zero at the 1% significance level in all six random effects models, with point estimates between 41.9% and 61.3%. The yr_{1986} dummy variable is statistically different from zero at the 5% significance level for only one of the random effects model, the point estimate is 30.3%. The yr_{1987} dummy variable is statistically different from zero at the 5% significance level in three of the random effects model, with point estimates between 33.8% and 36.7%. The $post_{1987}$ dummy variable is statistically different from zero at the 5% significance level in five of the random effects models, with point estimates between 36.9% and 59.3%.

The comparison group for both the starter and stopper dummy variables are long-relievers. The starter dummy variable is statistically different from zero at the 5% significance level for only one of the random effects model, the point estimate is 21.7%. The stopper dummy variable is statistically different from zero at the 5% significance level for one of the random effects models, the point estimate is 25.8%.

The start ratio variable is statistically different from zero at the 5% significance level in only one random effects model, the point estimate is -2.55. The quadratic start ratio variable is statistically different from zero at the 5% significance level in two random effects models, with point estimates between 2.56 and 4.01. The close ratio variable is statistically different from zero at the 5% significance level in only one

Table 6.1: Regression Table

	(1) ln(salary$_R$)	(2) ln(salary$_R$)	(3) ln(salary$_R$)	(4) ln(salary$_R$)	(5) ln(salary$_R$)	(6) ln(salary$_R$)
yr_{1984}	0.139	0.172	0.219	0.185	0.158	0.221
	(0.95)	(1.23)	(1.80)	(1.32)	(1.13)	(1.70)
yr_{1985}	0.540***	0.570***	0.613***	0.444***	0.419***	0.483***
	(4.09)	(4.64)	(5.94)	(4.03)	(3.73)	(4.60)
yr_{1986}	0.0512	0.0579	0.141	0.248*	0.243*	0.303**
	(0.31)	(0.36)	(0.91)	(2.09)	(1.98)	(2.70)
yr_{1987}	-0.0617	-0.0449	0.0252	0.359**	0.338**	0.367***
	(-0.26)	(-0.19)	(0.11)	(3.17)	(2.83)	(3.47)
$post_{1987}$	0.319*	0.372**	0.369***	0.564***	0.524***	0.593***
	(2.11)	(2.65)	(3.40)	(5.21)	(4.58)	(5.68)
starter	0.113			0.217**		
	(0.88)			(3.18)		
stopper	0.290*			0.258***		
	(2.44)			(4.70)		
ln(IPOuts$_{ma2l}$)	0.577***	0.587***	0.504***	0.585***	0.564***	0.563***
	(6.43)	(6.31)	(5.69)	(12.50)	(11.26)	(12.09)
ln(attendance$_{ma2l}$)	0.337*	0.272	0.432**	0.358***	0.323***	0.394***
	(2.24)	(1.87)	(3.18)	(5.24)	(4.63)	(6.10)
arbitration	0.0344	0.123	-0.0143	0.440***	0.456***	0.425***
	(0.21)	(0.79)	(-0.09)	(5.32)	(5.50)	(5.32)
free_agency	0.0385	0.0634	0.0121	0.430***	0.410***	0.410***
	(0.31)	(0.49)	(0.11)	(7.75)	(7.24)	(8.00)
left_throw	0.291	0.309	0.452	0.0524	0.0309	0.0778
	(0.82)	(0.84)	(1.15)	(0.82)	(0.48)	(1.31)
ln(strikeout_to_walk$_{ma2l}$)	0.307*	0.306*	0.0911	0.469***	0.504***	0.323***
	(2.41)	(2.21)	(0.61)	(7.14)	(7.62)	(4.46)
age	0.556***	0.470**	0.482**	0.457***	0.485***	0.430***
	(3.78)	(2.85)	(2.76)	(5.91)	(6.09)	(5.84)
(age)^2	-0.00688**	-0.00576*	-0.00593*	-0.00611***	-0.00649***	-0.00578***
	(-2.97)	(-2.28)	(-2.20)	(-5.17)	(-5.33)	(-5.17)
start_ratio$_{ma2l}$		-1.653*	-3.985*		0.122	-2.558***
		(-2.49)	(-2.56)		(1.30)	(-4.12)
(start_ratio$_{ma2l}$)^2		1.522*	4.047**			2.564***
		(2.16)	(2.92)			(4.96)
inter_closing_starting			7.552*			4.933***
			(2.48)			(3.99)
closing_ratio$_{ma2l}$			-1.680			-2.018**
			(-1.14)			(-3.11)
(closing_ratio$_{ma2l}$)^2			2.663			3.001***
			(1.94)			(5.09)
Constant	-6.368*	-3.710	-5.556	-4.929***	-4.651**	-4.375**
	(-2.19)	(-1.29)	(-1.79)	(-3.44)	(-3.17)	(-3.08)
Observations	281	281	281	1962	1962	1962

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Models 1-6 use a random effects estimator and clustered standard errors.
Models 1-3 are estimated using the small sample and 4-6 are estimated using the large sample.

random effects model, the point estimate is -2.02. The quadratic close ratio variable is statistically different from zero at the 5% significance level in only one random effects models, the point estimate is 3.00. The interaction term for the start and close ratio variables is statistically different from zero at the 5% significance level in only one random effects model, the point estimate is 4.93. The logic behind the quadratic start and close ratio terms is that there are three major categories of pitchers: starters, relievers, and closers. Historically, starting and closing pitchers are believed to be the most valuable, which fits the convexity of the start and close ratios-salary relationship.

The logged moving average strikeout-to-walk ratio variable is statistically different from zero at the 5% significance level in three of the random effects models, with point estimates between 32.3% and 50.4%. Thus, a one percent increase in the moving average of the strikeout-to-walk ratio, the regression estimates between a 32.3%-50.4% increase in a pitcher's salary, on average. The logged moving average of the outs pitched variable ($ln(ipouts_{ma2l})$), as defined by Lahman database, is statistically different than zero at the 5% significance level in all six of the random effects models, with point estimates between 50.4% and 58.7%. Thus, a one percent increase in the moving average of outs pitched, the regression estimates between a 50.4% and 58.7% increase in a pitcher's salary, on average.

The logged moving average attendance variable is statistically different from zero at the 5% significance level in four random effects models, with point estimates between 32.3% and 43.2%. Thus, a one percent increase in the moving average of attendance, the regression estimates between a 32.3% and 43.2% increase in a pitcher's salary, on average.

The arbitration dummy variable is statistically different from zero at the 5% significance level in three random effects models, with point estimates between 42.5% and 45.6%. The free agency dummy variable is statistically different from zero at the 5% significance level in three random effects models, with point estimates between

41.0% and 43.0%. The left-handed dummy variable is not statistically different from zero in any of the models. This is a confounding result, because it is well documented in the literature that left-handedness for pitchers is more valuable in Major League Baseball. This may be to blame for an increase in the use of left-handed pitchers, which would then diminish their advantage as compared to right-handed pitchers. Thus, eliminating any premium for being a left-handed pitcher.

The age variable is statistically different from zero at the 5% significance level in all six of the random effects models, with point estimates between 43.0% and 55.6%. The quadratic age variable is statistically different from zero at the 5% significance level in four of the random effects models, with point estimates between -0.7% and -0.6%. The concavity of the age function matches the belief that as experience increases wages increase, at a decreasing rate. The inflection point concerning age is when the risk of injury and reduced-reflexes and other physical qualities such as eyesight outweighs the gain in experience.

6.1 Subsection: Confidence Intervals

The small sample 95% confidence intervals for the year dummy variables are listed below: The yr_{1984} dummy variable is not statistically different from zero in any of the models. The yr_{1985} dummy variable's lower limit is 33.0% and the upper limit is 82.1%. The yr_{1986} dummy variable is not statistically different from zero in any of the models. The yr_{1987} dummy variable is not statistically different from zero in any of the models. The $post_{1987}$ dummy variable's lower limit is 2.3% and the upper limit is 64.8%.

The large sample 95% confidence intervals for the year dummy variables are listed below: The yr_{1984} dummy variable is not statistically different from zero in any of the models. The yr_{1985} dummy variable's lower limit is 22.8% and the upper limit is 68.9%. The yr_{1986} dummy variable's lower limit is 1.5% and the upper limit is 52.2%. The yr_{1987} dummy variable's lower limit is 13.6% and the upper limit is 58.3%. The

Table 6.2: Confidence Intervals (Small Sample):

	$ln(salary_R)$	**Coef.**	**Std. Err.**	**z**	**P**	**[95% Conf.**	**Interval]**
Model 1	age	0.556	0.147	3.780	0.000	0.268	0.844
Model 2	age	0.470	0.165	2.850	0.004	0.147	0.793
Model 3	age	0.482	0.175	2.760	0.006	0.140	0.824
Model 1	$(age)^2$	-0.007	0.002	-2.970	0.003	-0.011	-0.002
Model 2	$(age)^2$	-0.006	0.003	-2.280	0.023	-0.011	-0.001
Model 3	$(age)^2$	-0.006	0.003	-2.200	0.028	-0.011	-0.001
Model 1	arbitration	0.034	0.165	0.210	0.835	-0.289	0.358
Model 2	arbitration	0.123	0.156	0.790	0.430	-0.183	0.429
Model 3	arbitration	-0.014	0.159	-0.090	0.928	-0.325	0.296
Model 3	$closing\_ratio_{ma2l}$	-1.680	1.474	-1.140	0.255	-4.570	1.210
Model 3	$(closing\_ratio_{ma2l})^2$	2.663	1.370	1.940	0.052	-0.023	5.349
Model 1	free_agency	0.039	0.124	0.310	0.756	-0.205	0.282
Model 2	free_agency	0.063	0.130	0.490	0.626	-0.191	0.318
Model 3	free_agency	0.012	0.113	0.110	0.914	-0.209	0.233
Model 3	inter_closing_starting	7.552	3.046	2.480	0.013	1.583	13.521
Model 1	intercept	-6.368	2.914	-2.190	0.029	-12.080	-0.656
Model 2	intercept	-3.710	2.868	-1.290	0.196	-9.332	1.911
Model 3	intercept	-5.556	3.101	-1.790	0.073	-11.633	0.521
Model 1	left_throw	0.291	0.353	0.820	0.410	-0.401	0.983
Model 2	left_throw	0.309	0.367	0.840	0.400	-0.410	1.028
Model 3	left_throw	0.452	0.394	1.150	0.252	-0.321	1.224
Model 1	$ln(attendance_{ma2l})$	0.337	0.151	2.240	0.025	0.042	0.633
Model 2	$ln(attendance_{ma2l})$	0.272	0.145	1.870	0.061	-0.013	0.557
Model 3	$ln(attendance_{ma2l})$	0.432	0.136	3.180	0.001	0.166	0.699
Model 1	$ln(ipouts_{ma2l})$	0.577	0.090	6.430	0.000	0.401	0.752
Model 2	$ln(ipouts_{ma2l})$	0.587	0.093	6.310	0.000	0.404	0.769
Model 3	$ln(ipouts_{ma2l})$	0.504	0.089	5.690	0.000	0.330	0.678
Model 1	$ln(strikeout\_to\_walk_{ma2l})$	0.307	0.127	2.410	0.016	0.057	0.557
Model 2	$ln(strikeout\_to\_walk_{ma2l})$	0.306	0.139	2.210	0.027	0.034	0.578
Model 3	$ln(strikeout\_to\_walk_{ma2l})$	0.091	0.151	0.610	0.545	-0.204	0.386
Model 1	$post_{1987}$	0.319	0.151	2.110	0.035	0.023	0.614
Model 2	$post_{1987}$	0.372	0.141	2.650	0.008	0.097	0.648
Model 3	$post_{1987}$	0.369	0.109	3.400	0.001	0.156	0.582
Model 2	$start\_ratio_{ma2l}$	-1.653	0.664	-2.490	0.013	-2.954	-0.352
Model 3	$start\_ratio_{ma2l}$	-3.985	1.559	-2.560	0.011	-7.040	-0.930
Model 2	$(start\_ratio_{ma2l})^2$	1.522	0.705	2.160	0.031	0.141	2.904
Model 3	$(start\_ratio_{ma2l})^2$	4.047	1.387	2.920	0.004	1.329	6.766
Model 1	starter	0.113	0.128	0.880	0.378	-0.138	0.363
Model 1	stopper	0.290	0.119	2.440	0.015	0.057	0.523
Model 1	yr_{1984}	0.139	0.146	0.950	0.342	-0.147	0.426
Model 2	yr_{1984}	0.172	0.140	1.230	0.219	-0.103	0.447
Model 3	yr_{1984}	0.219	0.122	1.800	0.071	-0.019	0.458
Model 1	yr_{1985}	0.540	0.132	4.090	0.000	0.281	0.798
Model 2	yr_{1985}	0.570	0.123	4.640	0.000	0.329	0.811
Model 3	yr_{1985}	0.613	0.103	5.940	0.000	0.411	0.815
Model 1	yr_{1986}	0.051	0.163	0.310	0.753	-0.268	0.370
Model 2	yr_{1986}	0.058	0.160	0.360	0.717	-0.255	0.371
Model 3	yr_{1986}	0.141	0.155	0.910	0.363	-0.163	0.445
Model 1	yr_{1987}	-0.062	0.238	-0.260	0.796	-0.529	0.405
Model 2	yr_{1987}	-0.045	0.237	-0.190	0.850	-0.510	0.420
Model 3	yr_{1987}	0.025	0.219	0.110	0.909	-0.404	0.455

Table 6.3: Confidence Intervals (Large Sample):

	$ln(salary_R)$	**Coef.**	**Std. Err.**	**z**	**P**	**[95% Conf.**	**Interval]**
Model 1	age	0.457	0.077	5.910	0.000	0.305	0.609
Model 2	age	0.453	0.077	5.840	0.000	0.301	0.604
Model 3	age	0.430	0.074	5.840	0.000	0.286	0.574
Model 1	$(age)^2$	-0.006	0.001	-5.170	0.000	-0.008	-0.004
Model 2	$(age)^2$	-0.006	0.001	-5.170	0.000	-0.008	-0.004
Model 3	$(age)^2$	-0.006	0.001	-5.170	0.000	-0.008	-0.004
Model 1	arbitration	0.440	0.083	5.320	0.000	0.278	0.602
Model 2	arbitration	0.464	0.081	5.710	0.000	0.304	0.623
Model 3	arbitration	0.425	0.080	5.320	0.000	0.268	0.581
Model 3	$closing\_ratio_{ma2l}$	-2.018	0.649	-3.110	0.002	-3.289	-0.746
Model 3	$(closing\_ratio_{ma2l})^2$	3.001	0.590	5.090	0.000	1.844	4.157
Model 1	free_agency	0.430	0.055	7.750	0.000	0.321	0.539
Model 2	free_agency	0.418	0.055	7.580	0.000	0.310	0.526
Model 3	free_agency	0.410	0.051	8.000	0.000	0.310	0.510
Model 3	inter_closing_starting	4.933	1.237	3.990	0.000	2.509	7.357
Model 1	intercept	-4.929	1.434	-3.440	0.001	-7.741	-2.118
Model 2	intercept	-4.099	1.464	-2.800	0.005	-6.968	-1.229
Model 3	intercept	-4.375	1.421	-3.080	0.002	-7.159	-1.591
Model 1	left_throw	0.052	0.064	0.820	0.410	-0.072	0.177
Model 2	left_throw	0.030	0.063	0.480	0.635	-0.093	0.152
Model 3	left_throw	0.078	0.059	1.310	0.189	-0.038	0.194
Model 1	$ln(attendance_{ma2l})$	0.358	0.068	5.240	0.000	0.224	0.492
Model 2	$ln(attendance_{ma2l})$	0.334	0.069	4.860	0.000	0.200	0.469
Model 3	$ln(attendance_{ma2l})$	0.394	0.065	6.100	0.000	0.267	0.520
Model 1	$ln(ipouts_{ma2l})$	0.585	0.047	12.500	0.000	0.493	0.677
Model 2	$ln(ipouts_{ma2l})$	0.563	0.049	11.440	0.000	0.467	0.659
Model 3	$ln(ipouts_{ma2l})$	0.563	0.047	12.090	0.000	0.471	0.654
Model 1	$ln(strikeout\_to\_walk_{ma2l})$	0.469	0.066	7.140	0.000	0.340	0.597
Model 2	$ln(strikeout\_to\_walk_{ma2l})$	0.440	0.068	6.520	0.000	0.308	0.573
Model 3	$ln(strikeout\_to\_walk_{ma2l})$	0.323	0.072	4.460	0.000	0.181	0.464
Model 1	$post_{1987}$	0.564	0.108	5.210	0.000	0.352	0.776
Model 2	$post_{1987}$	0.564	0.110	5.140	0.000	0.349	0.779
Model 3	$post_{1987}$	0.593	0.104	5.680	0.000	0.388	0.797
Model 2	$start\_ratio_{ma2l}$	-1.205	0.263	-4.580	0.000	-1.721	-0.690
Model 3	$start\_ratio_{ma2l}$	-2.558	0.621	-4.120	0.000	-3.775	-1.341
Model 2	$(start\_ratio_{ma2l})^2$	1.326	0.258	5.140	0.000	0.820	1.832
Model 3	$(start\_ratio_{ma2l})^2$	2.564	0.517	4.960	0.000	1.551	3.577
Model 1	starter	0.217	0.068	3.180	0.001	0.083	0.350
Model 1	stopper	0.258	0.055	4.700	0.000	0.151	0.366
Model 1	yr_{1984}	0.185	0.140	1.320	0.187	-0.089	0.458
Model 2	yr_{1984}	0.199	0.137	1.450	0.146	-0.069	0.467
Model 3	yr_{1984}	0.221	0.130	1.700	0.089	-0.034	0.475
Model 1	yr_{1985}	0.444	0.110	4.030	0.000	0.228	0.659
Model 2	yr_{1985}	0.461	0.110	4.190	0.000	0.245	0.677
Model 3	yr_{1985}	0.483	0.105	4.600	0.000	0.277	0.689
Model 1	yr_{1986}	0.248	0.119	2.090	0.037	0.015	0.482
Model 2	yr_{1986}	0.275	0.121	2.280	0.023	0.038	0.512
Model 3	yr_{1986}	0.303	0.112	2.700	0.007	0.083	0.522
Model 1	yr_{1987}	0.359	0.113	3.170	0.002	0.137	0.580
Model 2	yr_{1987}	0.360	0.114	3.150	0.002	0.136	0.583
Model 3	yr_{1987}	0.367	0.106	3.470	0.001	0.160	0.574

$post_{1987}$ dummy variable's lower limit is 34.9% and the upper limit is 79.7%.

Table 6.4: Inflection Points (Small Sample):

	$ln(salary_R)$	**Coef.**	**Std. Err.**	**z**	**P**	**[95% Conf.**	**Interval]**
Model 1	Age	40.418	3.448	11.720	0.000	33.661	47.176
Model 2	Age	40.802	4.169	9.790	0.000	32.632	48.972
Model 3	Age	40.654	4.142	9.810	0.000	32.535	48.773
Model 2	Start Ratio	0.543	0.071	7.630	0.000	0.404	0.682
Model 3	Start Ratio	0.492	0.053	9.310	0.000	0.389	0.596
Model 3	Closing Ratio	0.315	0.123	2.560	0.010	0.074	0.557

Table 6.5: Inflection Points (Large Sample):

	$ln(salary_R)$	**Coef.**	**Std. Err.**	**z**	**P**	**[95% Conf.**	**Interval]**
Model 1	Age	37.418	1.193	31.360	0.000	35.080	39.757
Model 2	Age	37.075	1.118	33.150	0.000	34.883	39.267
Model 3	Age	37.214	1.148	32.430	0.000	34.965	39.464
Model 2	Start Ratio	0.454	0.034	13.310	0.000	0.388	0.521
Model 3	Start Ratio	0.499	0.037	13.500	0.000	0.426	0.571
Model 3	Closing Ratio	0.336	0.047	7.180	0.000	0.244	0.428

6.2 Subsection: Inflection Points

The inflection point estimate for age of a pitcher from the small sample was between 40 and 41 years old. Both the linear and quadratic estimates for age (small sample) were statistically different from zero at the 5% significance level. In addition, the linear and quadratic age variables (small sample) were jointly statistically different from zero at the 5% significance level. The inflection point estimate for age of a pitcher from the large sample was approximately 37 years old. Both the linear and quadratic estimates for age (large sample) were statistically different from zero at the 5% significance level. In addition, the linear and quadratic age variables (large sample) were jointly statistically different from zero at the 5% significance level. Based on the models estimates both functions are concave.

The inflection point estimate for start ratio of a pitcher from the small sample was between 49.2% and 54.3%. Both the linear and quadratic estimates for start ratio (small sample) were statistically different from zero at the 5% significance level. In addition, the linear and quadratic start ratios (small sample) were jointly statistically different from zero at the 5% significance level. The inflection point estimate for start ratio of a pitcher from the large sample was between 45.4% and 49.9%. Both the linear and quadratic estimates for start ratio (large sample) were statistically different from zero at the 5% significance level. In addition, the linear and quadratic start ratios (large sample) were jointly statistically different from zero at the 5% significance level. Based on the models estimates both functions are convex.

The inflection point estimate for close ratio of a pitcher from the small sample was 31.5%. Both the linear and quadratic estimates for close ratio (small sample) were not statistically different from zero. However, the linear and quadratic close ratios (small sample) were jointly statistically different from zero at the 5% significance level. The inflection point estimate for close ratio of a pitcher from the large sample was 33.6%. Both the linear and quadratic estimates for start ratio (large sample) were statistically different from zero at the 5% significance level. In addition, the linear and quadratic close ratios (large sample) were jointly statistically different from zero at the 5% significance level. Based on the models estimates both functions are convex.

CHAPTER 7: Conclusion

7.1 Discussion:

In the case when transparency is applied to Major League Baseball salaries, in favor of players, the decrease in asymmetric information increased wages. It is perhaps surprising the speed at which this new information was incorporated into pitchers' salaries. Transparency had a near immediate effect in 1985, the large sample random effects (model 3) point estimate for the real change in wages during 1985 was 48.3% over the 1984 level. Although, we hypothesized the transparency effect would have influenced salaries in 1986, the year after transparency was introduced, it appears the biggest impact was the first year of transparency. Several narratives could explain the speed of response to new wage information. Marvin Miller encouraged players to openly discuss their salaries hoping to put upward pressure on wages. Furthermore, Major League Baseball pitchers are not like most workers. They frequently have lawyers and agents negotiating to maximize their salaries, which would typically only be privileged to the most elite of executives or professional athletes. In contrast, Mas (2014) studied city managers in California and estimates their salaries declining by approximately 8% after transparency. However, the study may not adequately account for the coinciding municipal debt bubble in California after the Great Recession. Nonetheless, the findings could suggest that not all wage markets would respond like Major League Baseball. Maybe only markets like professional sports and corporate executives have the ability to incorporate this type of information, with aid of compensation consulting firms, into salary negotiations.

It is well documented that agents and wage consulting firms have privy information about other players' salaries, which would also put upward pressure on wages for players who hired such agents and firms. Transparency correlates with a large increase in Major League Baseball salaries. However, the reason for the decline in 1986 salaries might be that the 1985 momentum overshot the equilibrium wages for players' on average. In 1986, wages fell to 30.3% over the 1984 level, where in 1985 salaries had been 48.3% higher than in 1984. Additionally, wages might have been more efficient than originally thought in the pre-1985 period. Less plausible, but the salaries of pitchers in this period might have been so close to their marginal revenues that the increase in wages were not sustainable. However, I believe the strongest argument explaining the decline in 1986 is collusion.

A complicating factor is this period overlaps with a period of team-owner collusion. Collusion is suspect during the 1985, 1986, and 1987 seasons. Collusion is thought to put downward pressure on wages, which could explain the subsequent decline in pitcher's salaries in 1986. The table 7.1 shows the total payrolls by year and team from 1982 to 1989. Note the increase in total payrolls in 1985 and the decrease in total payrolls in 1987.

Table 7.2 shows the total amount the courts awarded in November 1990 for the three years of collusion. The courts estimated that $280 million (1990 dollars) in wages were stolen from players. From the settlement amount, it can be interpolated that the total Major League Baseball player compensation should have been approximately $1.3 billion (1990 dollars) between 1985 and 1987, which calculates to an average decrease in wages of approximately 21.0%. When you adjust the estimated 1985 wage increase (large sample random effects 3) for the estimated 21.0% decrease due to collusion, the transparency effect increases to 69.3% over the 1984 level. However, the 1985 estimate might be subject to upward bias, due to the potential momentum and wage equilibrium arguments already set forth and a weak collusion argument that soon will be explained. That is why the 1986 estimate for wage increases might be less susceptible to bias when estimating the transparency effect. When using the

Table 7.1: Major League Baseball Payrolls by Team (Nominal)

	1982	**1983**	**1984**	**1985**	**1986**	**1987**	**1988**	**1989**
Atlanta Braves	5,237,300	8,690,500	10,071,725	14,771,382	17,940,286	13,520,393	12,634,667	9,065,334
Baltimore Orioles	6,063,950	7,632,625	9,005,100	12,371,429	12,196,745	12,913,466	11,370,404	10,944,499
Boston Red Sox	6,187,825	6,620,825	7,446,950	11,080,695	16,003,236	12,145,206	15,629,592	19,064,885
California Angels	10,585,075	9,745,825	10,785,775	11,559,593	14,254,758	12,035,499	12,382,388	15,362,333
Chicago Cubs	5,516,550	6,723,675	10,554,850	13,478,225	16,904,832	12,732,999	12,849,333	12,167,000
Chicago White Sox	6,191,825	7,277,850	11,182,025	9,849,689	10,099,510	7,872,800	7,736,952	8,981,094
Cincinnati Reds	5,088,300	5,976,700	6,725,475	9,258,848	11,785,036	8,536,500	9,998,833	12,257,000
Cleveland Indians	5,400,000	6,053,350	3,994,350	6,623,133	8,047,000	7,955,250	10,244,500	10,349,500
Detroit Tigers	4,353,350	6,597,475	9,283,300	10,850,643	12,254,047	12,557,881	15,597,071	14,147,763
Houston Astros	7,664,125	9,120,625	9,574,775	10,153,335	10,368,276	11,358,371	13,565,576	16,761,625
Kansas City Royals	6,452,275	7,749,050	7,279,000	11,754,512	13,996,417	12,148,384	14,058,873	17,101,047
Los Angeles Dodgers	5,408,300	7,213,875	7,913,250	11,970,412	15,471,276	12,100,987	16,412,515	21,147,506
Milwaukee Brewers	8,274,125	8,801,525	9,630,375	12,216,965	8,429,321	7,617,724	10,864,000	11,901,500
Minnesota Twins	1,683,375	2,449,500	4,300,600	7,238,667	9,815,200	11,671,956	13,444,800	14,303,000
Montreal Expos	7,479,800	8,833,925	9,213,925	10,195,246	11,937,394	7,625,552	9,452,333	15,141,222
New York Mets	6,588,475	11,599,675	7,073,800	11,013,714	13,597,780	14,102,214	15,452,714	21,300,878
New York Yankees	10,299,700	11,596,675	11,463,600	15,398,047	17,248,360	17,082,214	21,524,152	18,482,251
Oakland Athletics	6,658,375	6,670,375	9,550,675	10,008,823	9,933,388	11,928,250	11,628,083	17,722,999
Philadelphia Phillies	9,759,250	11,054,125	10,036,900	11,785,445	11,715,166	11,301,833	13,248,000	8,633,000
Pittsburgh Pirates	6,280,850	7,869,225	8,266,525	10,223,945	10,231,500	4,024,500	8,647,500	12,463,000
San Diego Padres	3,448,650	6,545,500	7,779,975	9,801,052	11,897,522	9,484,429	10,978,168	14,004,000
San Francisco Giants	4,960,950	6,205,100	7,053,300	7,777,945	8,682,000	9,535,000	12,332,000	17,255,083
Seattle Mariners	2,860,125	2,971,875	4,212,625	5,549,870	6,382,309	5,499,500	6,545,950	8,702,500
St. Louis Cardinals	5,938,325	6,484,825	7,272,150	10,441,639	9,481,677	11,615,000	14,027,500	16,077,333
Texas Rangers	4,660,600	4,521,200	6,177,025	8,101,222	6,892,218	6,342,718	7,105,500	10,831,781
Toronto Blue Jays	3,196,500	5,327,175	7,390,800	11,800,281	12,447,880	11,626,334	14,098,725	16,016,666
Total:	156,237,975	190,333,075	213,238,850	275,274,757	308,013,134	275,334,960	321,830,129	370,184,799

Table 7.2: Estimated Major League Baseball Payroll (1985-1987)

Total MLB Player Compensation 1985-1987 (1990 Dollars):	1,056,432,868
Total Damages (1990 Dollars):	280,000,000
	1,336,432,868
Percent of Wages Lost:	21.0%

Table 7.3: Major League Baseball Payrolls by Team (1990 Dollars)

	1982	1983	1984	1985	1986	1987	1988	1989
Atlanta Braves	7,431,079	11,889,457	13,224,699	18,733,752	21,901,553	16,268,243	14,611,223	10,016,034
Baltimore Orioles	8,603,993	10,442,181	11,824,165	15,690,020	14,889,822	15,537,965	13,149,180	12,092,270
Boston Red Sox	8,779,756	9,057,939	9,778,233	14,053,052	19,536,797	14,613,566	18,074,671	21,064,258
California Angels	15,018,908	13,333,246	14,162,283	14,660,412	17,402,250	14,481,563	14,319,478	16,973,412
Chicago Cubs	7,827,300	9,198,648	13,859,067	17,093,710	20,637,468	15,320,821	14,859,471	13,442,978
Chicago White Sox	8,785,431	9,956,813	14,682,580	12,491,833	12,329,511	9,472,847	8,947,314	9,922,959
Cincinnati Reds	7,219,666	8,176,713	8,830,898	11,742,501	14,387,206	10,271,436	11,563,041	13,542,416
Cleveland Indians	7,661,930	8,281,577	5,244,789	8,399,765	9,823,801	9,572,054	11,847,140	11,434,873
Detroit Tigers	6,176,863	9,025,994	12,189,456	13,761,289	14,959,776	15,110,112	18,037,063	15,631,467
Houston Astros	10,874,442	12,477,911	12,572,177	12,876,931	12,657,621	13,666,817	15,687,763	18,519,450
Kansas City Royals	9,154,977	10,601,461	9,557,706	14,907,618	17,086,866	14,617,390	16,258,230	18,894,468
Los Angeles Dodgers	7,673,707	9,869,289	10,390,509	15,181,432	18,887,379	14,560,360	18,980,074	23,365,287
Milwaukee Brewers	11,739,957	12,041,351	12,645,183	15,494,122	10,290,540	9,165,930	12,563,554	13,149,634
Minnesota Twins	2,388,500	3,351,157	5,646,911	9,180,413	11,982,425	14,044,134	15,548,092	15,802,984
Montreal Expos	10,612,908	12,085,677	12,098,363	12,930,084	14,573,205	9,175,349	10,931,047	16,729,112
New York Mets	9,348,228	15,869,494	9,288,267	13,968,103	16,600,209	16,968,311	17,870,123	23,534,744
New York Yankees	14,613,996	15,865,390	15,052,303	19,528,518	21,056,848	20,553,959	24,891,371	20,420,522
Oakland Athletics	9,447,408	9,125,728	12,540,533	12,693,654	12,126,709	14,352,517	13,447,170	19,581,645
Philadelphia Phillies	13,847,165	15,123,129	13,178,972	14,946,848	14,301,909	13,598,788	15,320,505	9,538,360
Pittsburgh Pirates	8,911,747	10,765,873	10,854,377	12,966,481	12,490,645	4,842,429	10,000,307	13,770,020
San Diego Padres	4,893,206	8,954,887	10,215,512	12,430,149	14,524,529	11,412,020	12,695,583	15,472,627
San Francisco Giants	7,038,972	8,489,186	9,261,350	9,864,351	10,599,011	11,472,869	14,261,207	19,064,658
Seattle Mariners	4,058,162	4,065,817	5,531,396	7,038,603	7,791,541	6,617,204	7,569,993	9,615,149
St. Louis Cardinals	8,425,746	8,871,877	9,548,711	13,242,571	11,575,259	13,975,602	16,221,949	17,763,395
Texas Rangers	6,612,813	6,185,446	8,110,755	10,274,346	8,414,040	7,631,796	8,217,078	11,967,732
Toronto Blue Jays	4,535,437	7,288,099	9,704,505	14,965,664	15,196,408	13,989,240	16,304,317	17,696,366
Total (1990 Dollars):	221,682,297	260,394,337	279,993,698	349,116,220	376,023,327	331,293,321	372,176,945	409,006,819
Total Change (1990 Dollars):		38,712,040	19,599,361	69,122,521	26,907,108	(44,730,006)	40,883,623	36,829,875
% Change:		17.5%	7.5%	24.7%	7.7%	-11.9%	12.3%	9.9%

1986 change, the transparency effect decreased to 51.3% over the 1984 level, when adjusting for collusion.

Table 7.3 (above) takes the table 7.1 (From page 35) and converts all the amounts to 1990 dollars to allow for easier comparison. The second table then takes the amounts from the first table and calculates the percent change as compared to the previous period.

Collusion in its simplest form is when one team agrees not to bid up the contract of a player in exchange for the other team agreeing not to bid up the contract of another player. Therefore, collusion requires at least two teams to work properly. It should also be assumed that both teams should receive a comparable level of benefit or the they would be less inclined to collude. And if collusion was effective, you would assume both teams decreased their payroll for a given year. If you look at the table 7.4 (below), starting in the 1983 column, three teams, the Angels, Oakland and

Table 7.4: Major League Baseball Payrolls Percent Change by Team (1990 Dollars)

	1983	**1984**	**1985**	**1986**	**1987**	**1988**	**1989**
Atlanta Braves	60.0%	11.2%	41.7%	16.9%	-25.7%	-10.2%	-31.4%
Baltimore Orioles	21.4%	13.2%	32.7%	-5.1%	4.4%	-15.4%	-8.0%
Boston Red Sox	3.2%	8.0%	43.7%	39.0%	-25.2%	23.7%	16.5%
California Angels	-11.2%	6.2%	3.5%	18.7%	-16.8%	-1.1%	18.5%
Chicago Cubs	17.5%	50.7%	23.3%	20.7%	-25.8%	-3.0%	-9.5%
Chicago White Sox	13.3%	47.5%	-14.9%	-1.3%	-23.2%	-5.5%	10.9%
Cincinnati Reds	13.3%	8.0%	33.0%	22.5%	-28.6%	12.6%	17.1%
Cleveland Indians	8.1%	-36.7%	60.2%	17.0%	-2.6%	23.8%	-3.5%
Detroit Tigers	46.1%	35.0%	12.9%	8.7%	1.0%	19.4%	-13.3%
Houston Astros	14.7%	0.8%	2.4%	-1.7%	8.0%	14.8%	18.1%
Kansas City Royals	15.8%	-9.8%	56.0%	14.6%	-14.5%	11.2%	16.2%
Los Angeles Dodgers	28.6%	5.3%	46.1%	24.4%	-22.9%	30.4%	23.1%
Milwaukee Brewers	2.6%	5.0%	22.5%	-33.6%	-10.9%	37.1%	4.7%
Minnesota Twins	40.3%	68.5%	62.6%	30.5%	17.2%	10.7%	1.6%
Montreal Expos	13.9%	0.1%	6.9%	12.7%	-37.0%	19.1%	53.0%
New York Mets	69.8%	-41.5%	50.4%	18.8%	2.2%	5.3%	31.7%
New York Yankees	8.6%	-5.1%	29.7%	7.8%	-2.4%	21.1%	-18.0%
Oakland Athletics	-3.4%	37.4%	1.2%	-4.5%	18.4%	-6.3%	45.6%
Philadelphia Phillies	9.2%	-12.9%	13.4%	-4.3%	-4.9%	12.7%	-37.7%
Pittsburgh Pirates	20.8%	0.8%	19.5%	-3.7%	-61.2%	106.5%	37.7%
San Diego Padres	83.0%	14.1%	21.7%	16.8%	-21.4%	11.2%	21.9%
San Francisco Giants	20.6%	9.1%	6.5%	7.4%	8.2%	24.3%	33.7%
Seattle Mariners	0.2%	36.0%	27.2%	10.7%	-15.1%	14.4%	27.0%
St. Louis Cardinals	5.3%	7.6%	38.7%	-12.6%	20.7%	16.1%	9.5%
Texas Rangers	-6.5%	31.1%	26.7%	-18.1%	-9.3%	7.7%	45.6%
Toronto Blue Jays	60.7%	33.2%	54.2%	1.5%	-7.9%	16.5%	8.5%

Rangers, experience overall payroll declines as compared to 1982. None of which were similar in magnitude and only the Angels observe a substantial overall payroll decline of -11.2%. In 1984, five teams, the Indians, Royals, Mets, Yankees and Phillies, experience overall payroll declines as compared to 1983. However, in this case both the Indians and Mets observe the largest similar declines. The Indians observe a -36.7% decline, when the Mets observe a -41.5% decline. In addition, both the Royals and Phillies observe large similar declines. The Royals observe a -9.8% decline, when the Phillies observe a -12.9% decline. The two pairs of similar percent declines could suggest that collusion could have been taking place as early as 1984.

The 1985 transparency policy change might have masked collusion in the 1984-1985 period. In 1985 only one team, the White Sox experience overall payroll decline, which was -14.9% as compared to 1984. However, five teams experience very small payroll increases. Those teams were the Angels, Astros, Expos, Oakland, and Giants. The Angels observe a 3.5% increase in payroll. The Astros observe a 2.4% increase in payroll. The Expos observe a 6.9% increase in payroll. Oakland observes a 1.2% increase in payroll. Lastly, the Giants observe a 6.5% increase in payroll. Yet, the average increase in 1985 payrolls was 24.7%. This might suggest a weak form of collusion taking place in 1985, similar to the 1984 period.

In 1986, nine teams experience overall payroll declines as compared to 1985. The most notable decreases were the Brewers with a 33.6% decline, Cardinals with a 12.6% decline, and Rangers with a 18.1% decline. In 1987, 18 teams experience overall payroll declines as compared to 1986. The most notable decreases were the Braves with a 25.7% decline, Red Sox with a 25.2% decline, Cubs with a 25.8% decline, White Sox with a 23.2% decline, Reds with a 28.6% decline, Dodgers with a 22.9% decline, Expos with a 37.0% decline, Pirates with a 61.2% decline, and Padres with a 21.4% decline. In 1988, six teams, the Braves, Orioles, Angles, Cubs, White Sox, and Oakland observe overall payroll declines as compared to 1987. Lastly, in

Table 7.5: Large Sample T-Tests:

	yr1985 = yr1986		**yr1985 = yr1986 + 21.0%**	
	chi2(1)	**Prob >chi2**	**chi2(1)**	**Prob >chi2**
Model 1	5.050	2.5%	0.030	86.4%
Model 2	4.320	3.8%	0.070	78.9%
Model 3	3.980	4.6%	0.110	74.6%

1989, seven teams, the Braves, Orioles, Cubs, Indians, Tigers, Yankees, and Phillies observe overall payroll declines compared to 1988. The peak of collusion appears to have been in 1987, with 18 teams observing overall payroll declines. The most persistent potential colluders appear to have been the Braves, Orioles, Angels, Cubs, White Sox, Indians, Yankees, Oakland, Phillies, and Cardinals. All of which had three seasons with declines in overall payrolls, and the Phillies with four seasons. Only the Twins and Giants never observe a payroll decline during the 1983-1989 period.

The year 1985 was unlike the possible 1984 collusion because the teams were first subject to transparency, which might have blindsided many teams by how effectively it increased overall payrolls. It is also worth mentioning that the collusion-adjusted 1986 transparency estimate of 51.3% is only slightly higher than the 1985 non-collusion-adjusted transparency estimate of 48.3%. This compliments the idea that the 1984 and 1985 collusion was unlike that of the post-1985 period. Table 7.5 show t-tests for various large sample model estimates and their associated probabilities. The probability that the yr_{1985} estimate is equal to the yr_{1986} estimate is between 2.5% and 4.6%. However, when you add the collusion estimate of 21.0% to the yr_{1986} estimate the probability increases to between 74.6% and 86.4%.

It is impossible to know with certainty which teams or how many teams were involved in the collusion, but the possible 1984 and 1985 period of collusion does not appear to have been as dramatic or widespread as the post-1985 period of collusion. A possible reason for this is that teams had more incentive to collude after the 1985 spike in payrolls to combat the transparency effect. There does not appear to be any other reason for the declines observed in 1986 nor 1987; in fact, both years set new records for the highest attendance in a Major League Baseball season.[1] In the beginning Major League Baseball had a conventional wage market, where players, teams, and owners had minimal restrictions on wage contracts. However, that all changed on December 6, 1879 when Major League Baseball players became subject to the Reserve Clause. The Reserve Clause gave teams and owners nearly perfect monopsony power over their players. This marked the beginning of weakest period for a player's ability to freely negotiate their salaries. Since then Major League Baseball, with encouragement from players and the Major League Baseball Players Association has attempted to move back toward a convectional wage market. One key turning point was on February 28, 1968 when the first Collective Bargaining Agreement was finalized, which included the option for players to receive arbitration. The next major turning point, also in result of another Collective Bargaining Agreement, was free agency on December 23, 1975. Free agency marked the closest Major League Baseball has ever been to a conventional wage market since 1879. However, even with arbitration and free agency Major League Baseball is still radically different than many other wage markets. For example, most employees can choose at anytime to seek an alternative job, but this is not the case in professional baseball. Only after a player becomes a free agent can he then seek an alternative jobs as a baseball player. As a result of attempting to mitigate this gap between the differing wage markets, Major League Baseball and the Major League Baseball Players Association created this unique agreement allowing all players to know other players salaries.

In the introduction, the transparency thought experiment worked through how transparency can put upward pressure on wages. Yet, it also assumes the party not benefiting from transparency does not respond with a counter strategy. However, this is an unrealistic assumption. As illustrated by Major League Baseball, after

[1] http://www.baseballchronology.com/Baseball/Teams/Background/Attendance/

transparency was implemented in 1985, teams overwhelmingly began colluding as outlined by the number of team payrolls that observed declines after 1985. The team owners' counter strategy was collusion. However, that was not a sustainable strategy, because too many teams took part and the magnitude at which teams tried to collude was too large to go unnoticed as compared to the collusion in 1984. The backlash from players and the Major League Baseball Players Association combated collusion. Culminating in a court ruling in November 1990 requiring owners to pay back $280 million. Transparency created a price change in favor of the party receiving the asymmetric information.

REFERENCES

[1] G. Stone and L. J. Pantuosco, “Estimating baseball salary equations from 1961-2005: A look at changes in major league compensation,” *International Journal of Sport Finance*, vol. 3, no. 4, p. 228, 2008.

[2] L. M. Kahn, “Free agency, long-term contracts and compensation in major league baseball: Estimates from panel data,” *The Review of Economics and Statistics*, pp. 157–164, 1993.

[3] P. M. Sommers and N. Quinton, “Pay and performance in major league baseball: The case of the first family of free agents,” *The Journal of Human Resources*, vol. 17, no. 3, pp. 426–436, 1982.

[4] A. C. Krautmann and M. Oppenheimer, “Contract length and the return to performance in major league baseball,” *Journal of Sports Economics*, vol. 3, no. 1, pp. 6–17, 2002.

[5] J. Maxcy, “Motivating long-term employment contracts: risk management in major league baseball,” *Managerial and Decision Economics*, vol. 25, no. 2, pp. 109–120, 2004.

[6] A. C. Krautmann, E. Gustafson, and L. Hadley, “A note on the structural stability of salary equations: major league baseball pitchers,” *Journal of Sports Economics*, vol. 4, no. 1, pp. 56–63, 2003.

[7] C. Schmidt, “Does transparency increase executive compensation,” tech. rep., Working paper, Norwegian School of Economics, 2012.

[8] A. Mas, “Does transparency lead to pay compression?,” tech. rep., National Bureau of Economic Research, 2014.